Summer 2018

A Program Evaluation of A Technology Based Formative Assessment for Algebra Readiness

Cassandra Harris
College of William and Mary - School of Education, pmharrisc@gmail.com

A PROGRAM EVALUATION OF A TECHNOLOGY BASED FORMATIVE ASSESSMENT FOR ALGEBRA READINESS

A Dissertation

Presented to

The Faculty of the School of Education at

The College of William and Mary in Virginia

In Partial Fulfillment

Of the Requirements for the Degree of

Doctor of Education

---

by

Cassandra Harris

September 2017

A PROGRAM EVALUATION OF A TECHNOLOGY BASED FORMATIVE ASSESSMENT FOR ALGEBRA READINESS

By

Cassandra Harris

APPROVED on September 12, 2017

Margaret Constantino, Ph. D.
Chairperson of Doctoral Committee

Michael DiPaola, Ed. D.
Committee Member

Jacob Joseph, Ph. D.
Committee Member

Dedication

I dedicate this work to my son, Ryan Harris. Your support and confidence in me have urged me to finish well. I pray that the experience of watching me reach this dream of mine has taught you early on not to hesitate to reach for the stars and expect to land among them. God will send angels to assist you along the way. Remember, my precious son, that you are first a gentleman and then a scholar. With both come great responsibility to enrich the lives of others through your example and through your service. In all you do, you have a greatness about you that will stand out. When you enter a room, the anointing precedes you. Be comfortable in it; wear it well. It is your armor.

All my love,

Mom

Table of Contents

Acknowledgements

I have always said that God takes great pleasure in reaching far back into our wildest dreams to bring them into our reality, so I thank God for hand-selecting me as a cohort member in the 2014 Education Policy, Planning, and Leadership program. I would also like to thank all of the members of cohort 2014 for their impeccable models of leadership, which made this experience all the more enriching.

I wish to thank my dissertation chair, Dr. Peggie Constantino. In you, a continued example of professionalism, intelligence, grace, and strength overwhelmingly propelled me to keep moving on during the process of writing this dissertation. I would also like to thank the other members of my dissertation committee, Dr. Michael DiPaola and Dr. Jacob Joseph. Both have supported my efforts to make this final product something that I can honestly say I am very proud to have pursued. Also, Davis Clement, M.Ed., thank you for being my calm in the storm during this project.

Finally, I would also like to thank Kavansa Gardner, M.B.A., for your early recognition and ongoing support of my passion for the field of education and each prayer you've prayed for me along this journey.

LIST OF TABLES

LIST OF FIGURES

Abstract

Algebra 1 is often called the “gateway” course to higher education and opportunity. In the state of Virginia, the Algebra Readiness Diagnostic Test (ARDT) is the recommended mathematics formative assessment selected to monitor progress of students at each grade level. This program evaluation sought to take a closer look at tasks that teachers practice in the formative assessment process which exceed the mandatory ARDT periods for assessment. Teachers felt strongly about the adverse impact of assessment overload, but they also stressed the need to have continual alternative assessments, such as memory recall practices, to make certain that elementary students retained the mathematical concepts that were taught throughout the year. Further, teachers expressed the need to reinvent ways to keep elementary students engaged in the learning process, and most spoke about the need to vary instruction practices and intervention choices with peer tutoring, computer based instruction, and ongoing feedback. ARDT was not used often outside of mandatory assessment sessions; teachers opted to use more user-friendly mathematics software containing animations and varying content delivery methods. Teachers were consistent in their expressed belief that elementary students work best in assessment environments that are not delivered in mere black and white font, but are lively and changing in font color and delivery methods. Teachers also stated that using more than one formative assessment was a necessity because one would not meet the many needs of diverse student populations while continuing to keep the interest of elementary students.

Key Terms: Formative Assessment, Algebra 1, Elementary Mathematics, peer-tutoring, frequent assessment, Computer-Based Instruction,

A PROGRAM EVALUATION OF A TECHNOLOGY BASED FORMATIVE

ASSESSMENT FOR ALGEBRA READINESS

CHAPTER 1

INTRODUCTION

Description of the Problem

The *Coleman Report* of 1966 warned of the harmful impact widespread socioeconomic and racial segregation would have on the academic progress of racial and ethnic minorities and the poor (Coleman et al., 1966). Unfortunately, inner-city public schools continue in modern day segregation, resulting in academic challenges that are too widespread to number (Darling-Hammond, 2015). Representative of such challenges, disproportionately higher numbers of students from racial and ethnic minority backgrounds or of low socioeconomic status (SES) are not reaching proficiency in the fundamental mathematics courses intended to prepare them for higher-level mathematics courses, specifically Algebra 1 (National Assessment of Educational Progress, 2015). Algebra 1 has often been described as the gateway to quality post-secondary education experiences and rewarding career paths in science, technology, engineering, and mathematics (STEM; National Mathematics Advisory Panel [NMAP], 2008; U.S. Department of Education [USDOE], 1997; Walker & Senger, 2007; Walston & McCarroll, 2010). Currently, the United States economy continues to shift into one that is predominantly STEM-driven, but without change, few students in urban public schools will be academically prepared to be active participants in that economy (Langdon, McKittrick, Beede, Khan, & Doms, 2011; Vilorio, 2014).

The significance of Algebra 1 is that achievement in this subject is a predictor of success in those college and university mathematics courses most associated with jobs that command desirable wages both locally and abroad. In 2010, STEM job growth was three times faster than non-STEM job growth, and forecasts for 2020 suggest 9 million additional STEM jobs will be added to the United States economic system (Vilorio, 2014). While the projected need for qualified STEM talent continues to flourish, preparation for post-secondary STEM education continues to be determined by the successful completion of Algebra 1 (Schmidt, McKnight, Cogan, Jakwerth, & Houang, 1999; Snipes & Finklestein, 2015).

Background

Unfortunately for some students from racial and ethnic minority and/or low SES backgrounds, academic progress toward the completion of algebra can be hard to achieve. Though governmental policies such at the No Child Left Behind Act of 2001 (NCLB) are well intended, they have not produced the intended academic outcomes (DeAngelis, White, & Presley, 2010). One hurdle for public schools is the absence of Algebra 1 as a course option in some middle and high school curricula (Heppen et al., 2011). With or without Algebra 1 as a course offering, students who finish high school without having had Algebra 1 are more likely to be under-employed or inadequately prepared to demonstrate required mathematical competency at the time of college application (Gaertner, Kim, DesJardins, & McClarity, 2014). Further, in K-12 schools where Algebra 1 is available, students from racial and ethnic minority and/or low SES backgrounds are often underprepared and underrepresented in such courses (Domina, 2014; Stein, Kaufman, Sherman & Hillen, 2011; Stone, 1998; Walston & McCarroll, 2010).

In the primary years, mathematical literacy achievement for disadvantaged students attending public schools often lags behind that of students from higher income homes (Schmidt et al., 1999). This trend continues throughout their public K-12 educational experience (Duncan & Magnuson, 2011; Riegle-Crumb, 2006). Commonly, results of standardized tests demonstrate a significant and negative relationship between poverty and Scholastic Aptitude Test (SAT) mathematics scores (College Board, 2013). By the time many low-SES students that have attended urban public schools enter college, they are more likely than students of higher SES backgrounds to require remedial mathematic classes in preparation for the course requirements of their selected college program of study (Bailey, Jenkins, & Leinbach 2005; Epper & Baker, 2009).

The adverse gap in academic opportunity, teacher quality, and curricula experienced by students from racial and ethnic minority and/or low SES backgrounds is well documented and has been measured in SAT, Standards of Learning (SOL), high school graduation rates, and other assessments (Camara, 2013; Flores, 2007; Stronge, 2010; Virginia Department of Education [VDOE], 2016a). For example, the American College Testing (ACT) test measures high school student college readiness. In a 2015 administration of the ACT, only 14% of the 252,566 African American students and 29% of the 229,920 Hispanic students taking the test were found to be prepared for college mathematics, while 44% of all other students were classified as college mathematics ready. Such test results indicate a troubling lack of preparedness for post-secondary academics (ACT, 2015; Camara, 2013; Flores, 2007).

Disproportionately high numbers of students from racial and ethnic minority and/or low SES backgrounds are insufficiently prepared for post-secondary education, as evidenced by their need for college remediation classes (Complete College America [CCA], 2016; Strong American Schools [SAS], 2008). A study conducted by CCA with National Governors Association Common Completion Metrics data from 33 states and approximately 10 million students showed that (a) 70% of African Americans attending community college require at least one remedial course, and (b) 40% of African Americans and 30% of Hispanics are enrolled in both remedial math and English during the freshman year of college (CCA, 2016). Not only is the investment required to fund remediation costly and time consuming for students, but many students required to take college remediation classes never actually obtain the degree sought (CCA, 2012; Schneider, 2010).

Without college completion, students are less likely to qualify for the STEM jobs and careers predicted to dominate the workforce in years to come. Further, these students jeopardize their chances to become gainfully employed citizens in their communities and adversely impact America's ability to regain its position as a global leader in developing influential STEM innovation (Langdon et al., 2011; Obama, 2009). In Virginia's public schools, for example, there are in excess of 512,000 students living at or near poverty level (Duncombe & Cassidy, 2016). Ultimately, large numbers of undereducated students—including large numbers from minority backgrounds that are swiftly becoming the majority—will have an impact on local community quality of life, as well as the nation's intellectual and economic strength (Suitts, 2015).

There are many proposed causes of student shortcomings in math readiness. One area that is often overlooked is the power of elementary teachers in at-risk settings to directly impact student mathematic aptitude (Cross, Woods, & Schweingruber, 2009).

Student preparation for algebra was once thought to have the greatest influence on early middle school achievement, but there is a strong association with early pre-kindergarten and elementary mathematics strength and post-middle school success in mathematics (Jacobs, Franke, Carpenter, Levi, & Battey, 2007). Attracting quality teachers with strong mathematics backgrounds can be a challenge, however, particularly for urban schools in low SES neighborhoods (Darling-Hammond, 2015). Overall, teachers who possess lower mathematic aptitude are more likely to teach in low SES elementary schools (Stronge, 2010) and such teachers are often directly associated with adverse feelings toward mathematics in the students they teach (Sloan, 2010).

To add to this crisis, students in elementary schools, where foundations in mathematics are cultivated, are taught by one teacher each day in multiple subjects. In general, elementary teachers determine how much time is spent on each subject each day. It stands to reason that a teacher with lower levels of mathematical ability may opt to spend less time on mathematics during class time (Peker & Ertekin, 2011; Sloan, 2010). Consequently, limited exposure to mathematics in elementary classrooms may develop in students a fragile and inadequate relationship with mathematics, beginning a descent into mediocre mathematics achievement for years to come (Wu, Barth, Amin, Malcarne, & Menon, 2012).

Given STEM job dominance now and the projected growth of STEM jobs for the future, identification of mathematics gaps in achievement between students from racial and ethnic minority and/or low SES backgrounds and students from more advantaged or majority backgrounds must be reformed, and interventions and remediation must change. The purpose of this study is to evaluate a technology-based formative assessment program intended to assist in intervention, remediation, and improvement of student outcomes in preparation for Algebra 1. The program selected for use is an online formative assessment, the Algebra Readiness Diagnostic Test or more commonly called the ARDT.

Program Description

In 1997, the state of Virginia added higher mathematics to the existing K-12 graduation requirements. In an effort to increase the mathematical rigor across school divisions, three credits of Algebra 1 or higher mathematics would be the new academic goal (VDOE, 2016b). The state of Virginia sought to identify a methodology that would allow the state to incrementally assess the readiness of its students for Algebra 1 in the grades preceding the Algebra 1 course year. The VDOE made the ARDT program available to Virginia districts as a no-cost, online, formative assessment designed to inform instructional practices in gauging student progress towards Algebra 1 readiness.

ARDT utilizes computer adaptive technology (CAT) that adjusts the difficulty of test items presented to students based on the students' answers to previous items in the assessment. As such, CAT provides a more personalized experience for students that decreases the possibility of answer sharing and facilitates the identification of specific instructional needs for each child. By 2007, nearly 96% of school districts in the state were utilizing the ARDT formative assessment program (Cates, 2005; Linacre, 2000; Wallinger, 2008).

The purpose of this section is to offer an overview of the context in which the program evaluation will be conducted. Information will include a description of the region for the study, the current state of mathematics preparedness in the district and in the elementary schools of focus. Finally, a summary of the justification of program need within the district is provided.

Approximately 40% of school-aged children eligible to attend schools in the district live in poverty, and 75% of the student population is eligible for free or reduced-price lunch (VDOE, 2016a). The overall population of the region is increasing, and the increases may be attributed to lower housing costs, low rates of unemployment, short commute times, and a lower-than-average cost of living (Zasky, 2016). The student population of the division is approximately 24,000, with 38 different documented spoken languages (VDOE, 2016a).

The district, as do many high need school districts, experiences challenges in the area of teacher staffing. Though a large majority of teachers hold teacher certification, many of the teachers in the district are only in their first to third year in teaching (VDOE, 2016a). Such an inexperienced workforce may lack the wisdom of experience in the classroom required in urban public schools (Stronge, 2010). To add to an inexperienced workforce, the district experiences a continually high turnover rate in teachers each year. Education experts believe that teaching quality is most commonly realized in classrooms where teacher experience reaches 3 to 5 years, or where teachers have had specialized training in the unique needs of the type of school they are in (Darling-Hammond, 2010; Stronge, 2010).

The Virginia SOL test is a method of determining student academic progress toward minimum learning expectations. Measures of mathematics achievement across the school district vary considerably. Overall, 62% of students in the district of context met the mathematics annual measurable objective (AMO). In order for a school to be classified by the VDOE as meeting the minimum expectation for mathematics aptitude, a school must attain an overall student pass rate of 70% or more. This means that 70% of the schools' student population in Grades 3 through 5 must meet or exceed the SOL benchmark score to be considered "meeting benchmark." The average pass rate was calculated using "all-students" and encompassed all student subgroups. Average mathematics scores for the elementary schools for academic years 2014-2015, 2015-2016, and 2016-2017 with 1- and 3-year averages are found in Table 1. Accountability ratings given by year were based on the previous academic year's test (VDOE, 2016a).

Table 1

Pass-rate by School for Mathematics Benchmark, Grades 3–5

School	2014-2015		2015-2016		2016-2017	
	1 YR	3 YR	1 YR	3 YR	1 YR	3 YR
School F	71	62	76	71	67	71
School R	47	51	54	50	75	59
School B	86	80	90	84	88	88

Note. Expected percentage of pass-rate for mathematics benchmarks is 70% or more

The ARDT measures student mathematics aptitude and communicates by numeric score the grade level at which a student is performing. The ARDT can classify student grade-level performance from Grade 3 up to Grade 8 and also on Algebra 1 levels. The ARDT program for Virginia students is administered to students in Grades 6-9 who meet the following criteria:

- Have been unsuccessful in previous interventions and/or remediation programs, and/or

- have had below-average performance in the previous year's mathematic program, and/or
- did not pass the mathematics SOL assessment (Wallinger, 2008, p. 5).

Though ARDT targets students in Grades 6-9, districts may elect to use the formative assessment for students in earlier grades. In elementary schools, the same free, state-approved tool used in secondary schools can provide a means to assess students in lower grades according to the grade-level content of the test. In the context of this study ARDT will be used in fifth grade.

The fundamental component of the ARDT is the grade-level assessment. This assessment identifies students performing below, at, or beyond grade-level expectations in mathematics. These grade-level assessments are administered to students each school year during the fall and spring semesters. Schools that have failed to meet SOL benchmark requirements must also take a mid-year or winter grade-level assessment. In addition, the grade-level assessments continue to be available to teachers throughout the year and can be administered at the teachers' discretion to inform instruction. The results of the grade-level assessment yield an overall proficiency score per student, as well as supporting scores for various algebra sub-topics (VDOE, 2012).

Content strand tests found in ARDT consist of random sets of 10 questions that can be assigned to students in various subtopics of algebra. Content strands provide practice in explicit skill areas that are identified in grade-level testing. The content strand tests are selected by the teacher, per student, based on strengths and weaknesses identified in the pre-test. Teachers, tutors and instructional specialists use this data to differentiate instruction in the form of managed interventions. After completion of the test, students are assigned an overall score. The teacher correlates the students' scores to the supplied chart of score ranges to determine if they fall below, at, or beyond expected grade-level achievement.

Appendix A provides an example of the grade-level and strand test report output that is used for analysis of student progress. Teachers utilizing ARDT and similar technology-based assessments have access to real-time student data that can facilitate immediate adjustment to classroom instruction. The interventions selected by schools are generally comprised of some combination of the following:

- strand test
- tutors
- peer assistance (peer tutoring)
- instructional technology programs

The algebra readiness coordinator (ARC) oversees the ARDT process for the district and is responsible for professional development and trainings in all of the schools. In addition, the ARC is also responsible for class observations, suggestive feedback on interventions, and the oversight of ARDT testing periods within the district. The purpose of the ARC is to ensure that all parties in the ARDT process have the knowledge to perform the tasks assigned to them and to student performance for increased math competence. The ARC may also recommend tutoring for students.

The role of the teacher and the principal is to review the results of the ARDT and make adjustments in the classroom and in instruction to meet identified students' needs. Tutors are available in some schools as an additional resource to provide individual and small group sessions in satisfaction of the requirement of an additional 2.5 hours of instruction. Principals are also influential in planning testing schedules, determining the frequency of testing, and providing additional resources for intervention.

The way in which principals respond to the data available after administering the ARDT (post-ARDT) varies. Some principals have regular meetings with all teachers to review the scores by grade for each classroom. In doing so, principals also seek grade-level similarities in mathematics strengths and weaknesses to provide teachers an opportunity to exchange what methods might be working in their classrooms. Teachers are encouraged to make adjustments to increase overall strengths of student progress. Some principals seek positive competition between teachers by having collaboration meetings more often to compare mathematic assessment data by classroom or school. In addition, principals assign assistant principals and other designees to conduct routine observations in mathematics classrooms to document and share result yielding instruction.

In exchange for receipt of VDOE-provided formative assessment (ARDT) and funding resources to districts, districts are required to continually report and meet requirements mandated in VDOE oversight of the program. The VDOE also requires that districts make early identification of students in Grades 6-9 who are struggling with Algebra 1 readiness. Districts are provided the following guidelines from VDOE: (a) student/teacher ratios for intervention services are not expected to exceed 10 to 1, (b) intervention services should provide 2.5 hours in mathematics in addition to regular classroom time, and (c) spring and fall formative assessment results must be made available to VDOE whether districts elect to use the free ARDT assessment or otherwise (Wallinger, 2008). Figure 1 is a logic model that depicts the ARDT program. The programmatic activities that occur in school—outside of the mandated VDOE requirements—are uniquely those encouraged by school leadership.

Overview of the evaluation approach. Given the importance of social accountability and fiscal control in public schools, it is vital to assess the effectiveness of an intervention program such as ARDT and provide feedback to stakeholders. The selected program evaluation model is based on Stufflebeam's (2007) (context-inputs-process-products (CIPP) model, which provides a comprehensive approach to assessing context, inputs, process, and products, including program interventions, resources, and outcomes. The CIPP model, which can be used for both formative and summative inquiry, was selected in this program evaluation as a measure to highlight components that could be considered by stakeholders as part of the continuous improvement process.

Acronyms of the Program Evaluation

AMO – Annual Measurable Objective

ARC – Algebra Readiness Coordinator

ARDT – Algebra Readiness Diagnostic Test

CBI – Computer Based Instruction

CIPP – Context, Inputs, Process, and Products

SOL – Standards of Learning

STEM – Science, Technology, Engineering, and Mathematics

VDOE – Virginia Department of Education

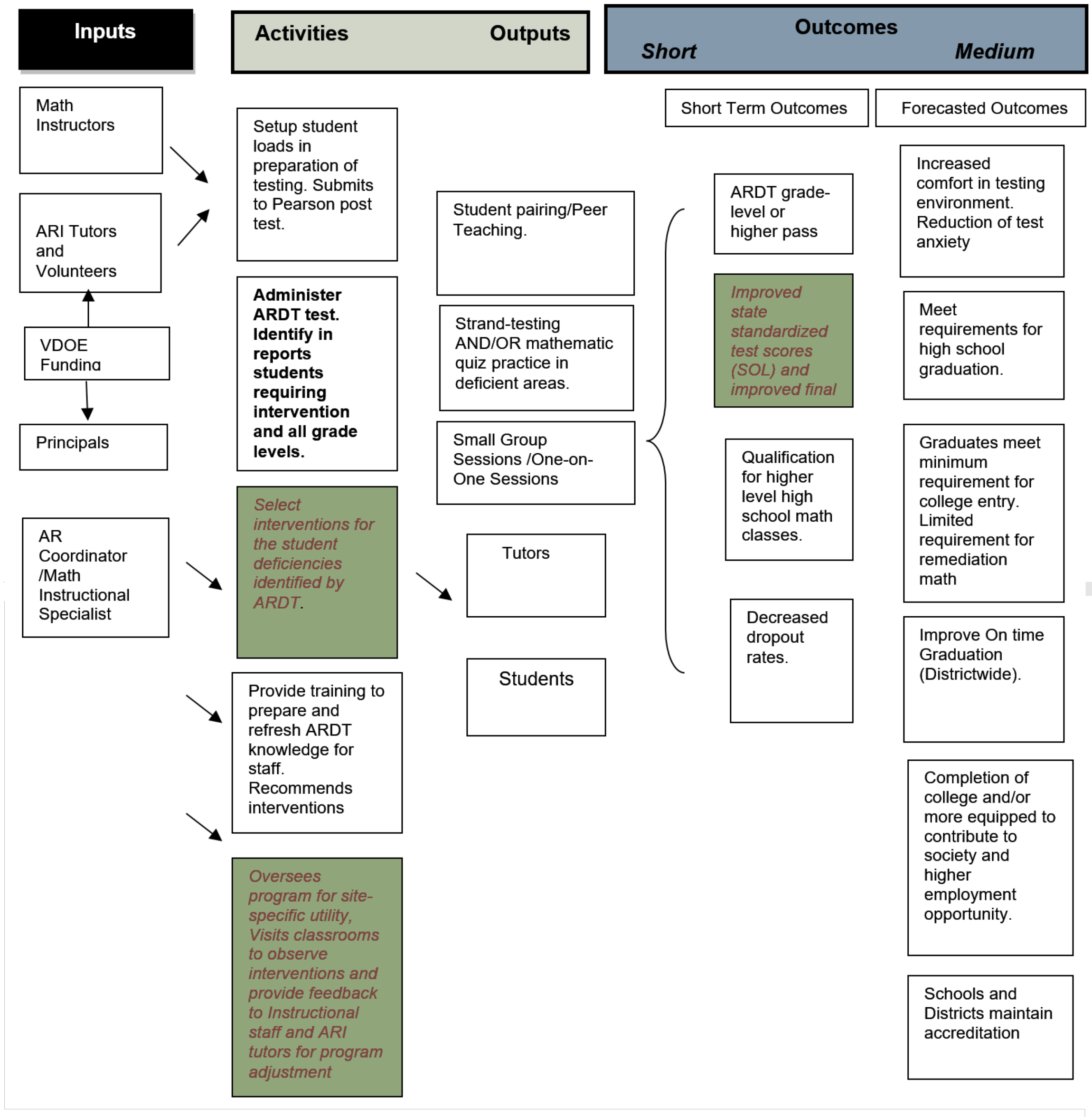

Figure 1. *Algebra Readiness Diagnostic Test (ARDT) Program Logic Model* utilizing the Stufflebeam (2007) CIPP model of evaluation (derived from Stufflebeam, D. L. (2007). *CIPP evaluation model checklist: A tool for applying the CIPP Model to assess long-term enterprises*). Inputs are comprised of district stakeholders and funding sources within the context of the student. Outputs are comprised of the processes and/or activities known to occur during the formative assessment process. Outcomes list the known results of the program that historically vary.

Program evaluation model. Since 1965, the U.S. federal government has employed the CIPP model as means of evaluation when confronting disparate education quality and conditions in poor urban and rural schools (Tseng, Diez, Lou, Tsai, & Tsai, 2010). The holistic information obtained by these schools through the CIPP component-based model encouraged feedback and analysis through stakeholder involvement and provided insight on opportunities for improvement. In comparison, the ARDT program, the subject of this evaluation, was initiated to provide resources for improving the results in a low SES public school struggling with preparing students for Algebra 1 (Stufflebeam, 2007; Tseng et al., 2010).

The selection of Stufflebeam's (2007) model for the evaluation approach is influenced by its apparent fit for evaluating the ARDT program and the historic use of the approach in education. In this study, process evaluation is intended to assess the implementation of strategies and interventions chosen by teachers based on formative assessments (Stufflebeam, 2007). Product evaluation will be used to assess whether the program is meeting the intended short-term outcome of improved algebra readiness.

Purpose of the evaluation. This study is aligned with the pragmatic paradigm, in which studies are conducted in authentic settings to solve problems of practice through mixed-methods research. The pragmatic approach encourages the use of both qualitative and quantitative data to inform program improvement (Stufflebeam, 2007). This program evaluation is representative of the use branch of evaluation, designed to focus on data that will be useful to stakeholders. This study is also intended to provide program stakeholders with information that can be used to improve student preparation for Algebra 1 (Mertens & Wilson, 2012).

Focus of the evaluation. This pragmatic program evaluation focuses on the process of interventions thought to directly contribute to the overall outcomes of the ARDT program. In doing so, this focus will assess the implementation of program plans that can be shared within and among schools to help stakeholders make adjustments that improve the program's effectiveness. The CIPP model in this context will be used to review teacher strategy selection. Finally, the overall outcomes or product—student mathematic outcomes by school—will be examined as it relates to the use of the ARDT, to determine intended and unintended outcomes.

Standards of program evaluation. Quality standards of program evaluation will be taken into account in developing this program evaluation. The Joint Committee on Standards for Educational Evaluation developed The Program Evaluation Standards to guide the evaluation of educational programs in a variety of contexts and provide a framework for defining the quality of evaluations (Yarbrough, Shulha, Hopson, & Caruthers, 2011). The Standards are organized into five categories: utility, feasibility, propriety, accuracy, and meta-evaluation (Yarbrough et al., 2011). The utility standards address the usefulness and appropriateness of the evaluation. The feasibility standards address the degree to which the evaluation can be done successfully in a given setting. The propriety standards ensure the fair, moral, ethical, and legal treatment of participants. The accuracy standards assess the dependability and trustworthiness of the evaluation. The meta-evaluation standards refer to a critical examination of the program evaluation itself to ensure the merit of the study (Mertens & Wilson, 2012).

Evaluation questions. The VDOE identifies student performance on the SOL assessments as one of the most reliable indicators of student progress (VDOE, 2016). The

ARDT program is provided to districts by the VDOE as an additional diagnostic tool to assess student strength in mathematics toward Algebra 1 success (VDOE, 2012). With student formative assessment data gleaned from the ARDT, teachers are better equipped to identify those students in need of interventions and those who may be ready for advanced placement. The reporting made available in ARDT offers a detailed accounting of both student weaknesses and strengths, enabling teachers and others in the program to assess individual needs and address them with differentiated instruction. The questions of the study are developed to address the varied interventions selected in the assessment process of ARDT and the beliefs that shaped the selection of those interventions used in algebra readiness planning. Data will be obtained through mixed methods research design to address the following evaluation questions:

1. Is there a difference in fifth-grade mathematics scores at the end of the year compared to scores at the beginning of the school year, as measured by reported ARDT scores?
2. Does the frequency of ARDT formative assessment administration effect the algebra readiness of students as measured by ARDT reporting scores?
3. How do teachers use ARDT grade-level progress reports to guide instructional intervention choices?

CHAPTER 2

REVIEW OF LITERATURE

Academic accountability requirements extended in the No Child Left Behind (NCLB) guidelines encouraged districts to look beyond the traditional teaching model to seek higher outcomes for students (Curry, Mwavita, Holter, & Harris, 2016; No Child Left Behind [NCLB], 2002). Rising English-language learning (ELL) populations, students with disabilities, and high populations of low SES students contribute to the need for varied pedagogies in classrooms so that all students progress academically. Though NCLB instituted end-of-year testing as an accountability measure for schools to demonstrate adequate yearly progress (AYP), students in poorer districts continue to lag behind their higher SES peers in mathematics (Lee, Grigg, & Dion, 2007). Educational leaders now view formative assessment as a means to intercept the awarding of poor grades at the end of the year in support of assessments that provide continual check points for possible intervention (Alvarez, Ananda, Walqui, Sato, & Rabinowitz, 2014; Marshall & Drummond, 2006; National Center on Educational Outcomes [NCEO], 2012; World-class Instructional Design and Assessment [WIDA], 2009).

Conceptual Framework

Formative assessment is not an event but a compilation of several actions requiring student and teacher interaction with an expectation of positive outcomes.

Wiliam and Thompson (2007) provided a framework for formative assessment identifying three processes:

- establishing where learners are in their learning,
- establishing where they are going, and
- establishing how to get there (Wiliam & Thompson, 2007).

This framework suggests a need for strategies that drive students and teachers to reach a defined goal, as well as promote two- or three-way communications throughout the process. Formative assessment is a continual process of identifying where students are in their learning, communicating and establishing the intended learning goal, and navigating each student using instruction or interventions that meet their individual needs in goal attainment. To do so, education institutions must not see assessment as a dated event in the calendar, but as part of a strategic, contiguously interconnected process (as depicted in Table 2).

Further, formative assessment is more considerate of student-teacher interaction, not restricted to merely capturing data. To support a successful formative assessment process, teachers are encouraged to relinquish ideas of themselves as single distributors of information, and to incorporate student feedback and interaction into the learning process. Moreover, the task of improving mathematical aptitude during progression to Algebra 1 requires a strategic approach for not only the teacher responsible for the mathematic development of young minds, but also for the students who must embrace new ways of approaching mathematical challenges.

Table 2

Key Strategies of Formative Assessment

Participants	Identification of Primary Goals	Establishing Current Status	Action Items
Teacher	Defining and documenting goals for learning	Designing active classroom discussions and activities that promote learning	Prioritizing feedback that will inform learning and applying change to instruction
Peer Student	Sharing learning opportunities	Promoting students as teachers and learners in the learning process	
Student	Comprehension of objectives for learning and measurements used for goal attainment	Endorsing students as the responsible parties for learning while providing support.	

Note. Based on Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. Educational Assessment, Evaluation, and Accountability 21(1), 5–31.

Education experts in effectual mathematics instruction highlight three skill-based components that are valued in supporting the promotion of growth in mathematical knowledge:

- conceptual understanding,
- procedural knowledge/fluency, and
- procedural flexibility (Rittle-Johnson & Star, 2007; Shellard & Moyer, 2002).

Conceptual understanding is knowledge of the various mathematical concepts that can be applied to problem solving but are not be restricted to one mathematical problem type. Moreover, conceptual understanding is knowledge of operations and relations in mathematics. Next, *procedural knowledge/fluency* refers to the learning of sequential actions required to solve mathematical problems. Grasping procedural knowledge, students are better able to reenact the actual steps required to reach correct problem resolution with higher consistency. Additionally, familiarization with multiple methods that can be used to resolve the same types of questions is considered *procedural flexibility*. Operating with procedural flexibility, students are more apt to accurately resolve unknown problems by using multiple methodologies to resolve problems. Early learning of multiple methods to resolve problems is thought to strengthen student mathematics aptitude with new problems and concepts, promoting agile problem resolution in students (Shellard & Moyer, 2002). Mathematics courses taken prior to Algebra 1 in environments that place emphasis on procedural flexibility, procedural knowledge, and conceptual knowledge help to build strong foundations needed to respond to the very different rigor requirements of Algebra 1 (Schiller, Schmidt, Muller, & Houang, 2010).

Algebra Readiness Initiative

The United States has long struggled with the notion of students' academic success being predicated on the incomes and zip codes of their guardians. The federal government has made strides to change such trends in mathematics through legislation to combat inequalities in the academic experiences of students. In the state of Virginia during 2001-2002, the Algebra Readiness Initiative (ARI) was instituted to offer support in the form of resources needed to address K-12 shortcomings in preparing students for Algebra 1. At the heart of ARI were the recommendations for districts to (a) use a VDOE-supplied standard formative assessment—the ARDT, and (b) accompany ARDT with intervention services for students demonstrating need (Wallinger, 2008).

The ARI in Virginia was further mainstreamed by the 1997 addition of Algebra 1 as a requirement for earning a standard diploma (Regional Educational Laboratory West

[REL], 2008). The VDOE approved and championed the use of their supplied formative assessment, the ARDT. Should a division elect not to use the VDOE-supplied ARDT to satisfy the requirement for a formative assessment, the division would be required to identify a formative assessment to be approved for use by the VDOE. Once approval of the alternative test was provided by VDOE, districts electing to use the alternative test would then need to submit pre- and post-year assessment reports to VDOE. Conversely, divisions electing to use the ARDT benefited from the reporting capabilities of the test. Additionally, the ARDT test is provided to districts at no cost making ARDT attractive for many struggling districts. More importantly, the ARDT test content avoids the issue of nonalignment with end-of-year SOL testing because both ARDT and SOL contents are developed and delivered by the same vendor, Pearson. By the year 2009, 90% of Virginia districts had elected to use ARDT (Wallinger, 2008).

The selection of approved services included financial assistance that could be used towards mathematics teachers, mathematics tutors, transportation costs to and from the sites of intervention, and other miscellaneous costs. The resources offered by VDOE did not come without requisite. The financial investments into Virginia school districts came with guidelines governing the receipt of funding assistance. The incentive funding from VDOE for intervention services were determined by a formula that considered the number of students that have been identified as requiring assistance and the composite index of the division's ability to pay. The research indicating students in poorer school districts were less likely to be taught by high quality teachers (Darling-Hammond & Sykes, 2003; Stronge, 2007) further supported the need for all teachers supported by ARI funding to be licensed by the Board of Education. Overall, VDOE offered a systematic plan to offer districts a means to prepare for its mathematic SOL testing while including resources to mitigate the effects of barriers associated with the adoption of formative assessment (VDOE, 2016a).

Best Practices in Algebra

Algebra 1 is among the first abstract reasoning mathematics courses students face in school (Vogel, 2008). The abstract nature of algebra is thought to be a contributor to the difficulty some students find in Algebra 1 (Carraher & Schliemann, 2007). Thus, the transition from a primarily numbers-driven mathematical setting to one that is filled with symbols requires students and teachers to participate in feedback exchange that will continually stimulate cognitive thinking.

Explicit instruction. One best-practice approach for algebra instruction is explicit instruction. In explicit instruction, teachers model mathematical solutions to problems for students, articulating and encouraging the thought process for problem resolution aloud, providing multiple examples of problems, and encouraging student-to-teacher feedback throughout the learning process (Hanover Research, 2016). Teachers demonstrate step-by-step resolutions of various types of problems, while explaining the thinking process of step selection. Additionally, as teachers encourage students to model such processes in mathematical problem resolution in the classroom, feedback from teachers can be immediate. In turn, this immediate feedback exchange facilitates effective problem solving and corrects errors in thinking prior to building upon them (Hinton, Stroizer, & Flores, 2015). Exposure to not only multiple types of problem but multiple methods of problem resolution has also been identified as optimal for algebra instruction (Hanover Research, 2016).

Multiple heuristics. The use of multiple heuristic strategies promotes the ability to conceptualize varied processes to solve problems in mathematics (Jayanthi, Gersten, & Baker, 2008). Heuristics are not problem-specific but are useful in organizing information to answer problems. The driving support for adopting multiple heuristics is to empower students with choices that will provide multiple tools to resolve many types of problems with fluency. One method used to stimulate use of multiple heuristics—common in countries such as Hong Kong and Japan, where students have historically performed well in mathematics—is the comparison method, which makes it common practice to present multiple ways to resolve problems (Stigler & Hiebert, 1999). In this method, the teacher can model problem resolution options by placing varied examples of methods to solve problems side-by-side, explaining each answer step by step. Comparing problem resolutions can be modeled by the teacher, but can also be a source of shared comparison between students to strengthen mathematic fluency (Smith, 2014).

Remediation Alternatives

The depth of success demonstrated in Algebra 1 prerequisite courses is a predictor of Algebra 1 course success (Schiller et al., 2010). For students struggling in mathematics prior to reaching Algebra 1, early bridge programs strengthen areas upon which algebra are founded and may help to increase later student pass rates (Schiller et al., 2010). Bridge programs help transition students into algebra by providing early exposure to subject matter they will encounter later.

Bridge programs. Algebra bridge programs can be offered during summer in preparation for fall algebra or during the normal school year in semester course offerings prior to Algebra 1. For lower-income student populations, summer learning losses especially in mathematics are more distinct (Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996; McCombs et al., 2011). Thus, bridge programs may meet a special needs requirement in lower income school districts during the summer transition from middle school to eighth or ninth grade, where Algebra 1 is usually taken. Although formal research on bridge programs is limited, the findings of research do suggest that these programs can be positively impactful for promoting stronger mathematical skill level in students approaching Algebra 1 (Herlihy & Quint, 2006; Snipes, Huang, Jaquet, & Finkelstein, 2016). Additionally, algebra bridge programs devote additional time to immersing students in algebraic foundational concepts in a brief instructional period. An alternative program that increases normal course time, double-dose algebra, is also thought to be effective as an Algebra 1 remediation alternative.

Double-dose algebra. Double dose algebra places at-risk students into an additional class period of algebra during school semesters and in effect doubles student time spent in algebra-related coursework (Hanover Research, 2016). Catholic schools, for example, have a long history of placing students who struggled in mathematics and were underprepared to go on to high school into double-dose settings (Bryk, Lee, & Holland, 1993). Results of the Catholic school double-dose use appeared promising as these programs both halved the number of students requiring remedial classes by 11th grade, and increased the number of students considered college ready (Bryk et al., 1993). There are other known applications of double-dose as an intervention strategy.

A well-known study of Chicago Public Schools (CPS) also reported promise in the use of double-dose methods. Chicago started double-dose implementation as a response to high freshman failure rates in algebra, approximately 25% (Durwood, Krone

& Mazzeo, 2010). However, there were advantages and disadvantages noted in the Chicago implementation of double-dose as a strategy. Students who were weaker in mathematics did improve their mathematics aptitude, and improved their mean grade point average (GPA) by .14 grade points on a 4.0 grading scale after their freshman year; however, immediate increases in GPA were not associated with double-dose investments (Hanover Research, 2016). Most research surrounding double-dose algebra points to the CPS implementation of double-dose Algebra 1. Similarly, a 2009 report by the Council of the Great City Schools noted that 49% of the 53 urban districts responding to their school survey identified double-dose Algebra 1 as their most widely used method for addressing weaker mathematic skills in students (Smith, 2014). Both double-dose and bridge programs add more time to algebra instruction, as well as keep students on schedule for completion of Algebra 1 prior to their freshman year in high school or shortly thereafter.

Formative Assessment

Academic gaps in mathematics aptitude begin to surface early in low SES and high minority population settings (Darling-Hammond, 2015); mathematics averages that fall below Virginia's expected scores are also representative of many of the elementary schools in the context of the study (VDOE, 2016a). In districts with more inexperienced and transient teacher workforces, district leadership seek ways to regularly view updates on student progress, and validate the accuracy of teachers' intervention selection. The process of obtaining continual feedback on student progress so that instruction can be adjusted accordingly is a highly recommended practice (NMAP, 2008). The ARDT program targeting mathematics aptitude was one of the first governmentally sponsored computer-based formative assessments in Virginia with a goal of early monitoring of student progress in mathematics through the transition to middle school. However, identification of student mathematics needs in fifth grade leaves little time for remediation of gaps prior to entering middle school, where mathematical concepts build upon concepts that were expected to be mastered on the elementary school level.

The more frequently teachers assess students, the earlier the teacher can identify student academic challenges and adjust instruction. A report released under the approval of the U.S. Department of Education (USDOE) specifically supports that elementary school students who are struggling in mathematics be assessed frequently (e.g., weekly or biweekly) so that instruction can be adjusted to student needs (Jayanthi et al., 2008; NMAP, 2008). Such recommendations for frequent formative evaluations are further supported by meta-analyses of 30 studies of no less than 3,835 people resulting in highly consistent and positive outcomes of formative evaluation programs that identify student academic needs ($d = .90$; Hattie, 2009, p. 181). In comparison, Hattie's (2009) reviews of meta-analyses demonstrated that, without formative assessment, teacher perception of student need had lower effect sizes, from $d = .25$ to $d = .40$. This supports the assertion that instruction without formative assessment could result in missed opportunities to differentiate instruction or intervene. Districts that begin monitoring formative assessments too late in elementary school allow weak foundations in mathematics to grow, likely unnoticed and unreported. In the context of this study, district-monitored and mandated elementary student mathematics formative assessment begins in fifth grade. District-supported formative assessment as an embedded practice throughout the life of elementary (i.e., third through fifth grades) may serve as a catalyst for building stronger mathematic skills early as has been realized in similar studies.

At-risk populations require elementary teachers who can identify student needs and respond, because each year of mathematical learning is cumulative and thought to strengthen foundations in mathematics for subsequent mathematics successes (Schiller et al., 2010). Teaching staff in inner city schools like the one in this study are often inexperienced, having three or fewer years of experience in the classroom (Stronge, 2010). In these schools, the consistent use of a formative assessment process is promising, since researchers like Martinez and Martinez (1992) have reported that formative assessment use produces even higher gains for inexperienced teachers. As a result, if teachers can identify and remediate problems that are diagnosed early in elementary school, future retention rates may also be reduced in secondary schools despite a historically inexperienced teacher workforce.

High student retention rates or repeating grade levels have been associated with increased student dropout (Andrew, 2014; Foster, 1993; Hattie, 2009); thus, reports associating formative assessment practices with achievement gains (Hattie & Timperley, 2007) also increase the appeal of formative assessment as a practice for proactively and actively improving student academic success throughout the school year. In the context of this study, Algebra 1 is a requirement for graduation. To increase first time pass rates, elementary school instruction must become more effective in strengthening mathematical skills in students earlier in elementary school. Properly monitoring student progress for strong mastery in foundational mathematics throughout elementary school will likely result in fewer dropouts in the district of context, because students may be much better prepared for secondary school mathematics and Algebra 1.

Though research findings strongly support the positive impact of formative assessment practice, the use of formative assessment in schools presents obstacles for some. One obstacle teachers voiced is a lack of time due to the periods of instruction required for standardized testing preparation (Abrams, 2007; Antoniou & James, 2014). Similarly, resources are more readily available to support summative assessment efforts. For example, the focus of national and local policy on summative assessment supports district allocations of calendar time to prepare and administer summative assessments. In addition, the sanctions of job loss and school accreditation loss associated with negative summative assessment outcomes implore school leaders to highly prioritize summative assessment (Antoniou & James, 2014). In the district of context, where accreditation status percentages fluctuate, each year there are elementary schools with low mathematics scores directly contributing to loss of accreditation. When school accreditation baselines are not met in summative scores, principals may face job loss or demotion. Not surprisingly, some principals and teachers may shy away from formative assessment exploration in favor of the familiarity and priority associated with summative assessments such as SOL tests (Abrams, 2007; Antoniou & James, 2014; Clark, 2011). Yet, research findings support that an emphasis on strengthening students throughout the year with formative assessment monitoring and planned intervention application has a positive impact on end-of-year summative assessment scores (Johnson & Kiviniemi, 2009; Peat, Franklin, Devlin, & Charles, 2004).

Interventions

In K-12 school divisions, the gathering and review of student formative assessment data are but part of the formative assessment process. While the review of formative assessment data can provide teachers a glimpse of student progress, the value of formative assessment data is in its utility to inform instruction. In large school settings where classrooms may contain high student-to-teacher ratios, formative assessment data can provide teachers information for early identification of interventions (Popham, 2008).

In the context of this program evaluation, principals and teachers use ARDT student data to guide which instructional interventions in the program will best meet the academic needs of students. The interventions used more in the context of the program evaluation are (a) testing frequency used to encourage forced memory recall (Gates, 1922), (b) computer-based instruction to reinforce and differentiate instruction, and (c) tutoring and peer tutoring.

Test frequency. The research on the impact of using smaller-scale tests more frequently as a part of the formative assessment process is promising (Gholami & Moghaddam, 2013). In the ARDT process, beyond required formative assessment periods, teachers may use formative assessment quizzes as a tool at their discretion. In doing so, teachers can periodically administer ARDT to assess student learning and receive reports on all areas of expected algebra readiness. Frequent administration of quizzes not only produces higher outcomes in summative assessments (Johnson & Kiviniemi, 2009; Peat et al., 2004), but also reduces the anxiety some students may experience in approaching high-stakes summative assessments (Dustin, 1971; Gholami & Moghaddam, 2013). Researchers refer to the occurrence of forced recall of information, as practiced in frequent quiz administration, as the *testing effect*, a phenomenon that has been studied dating back to the early 1900s (Gates, 1922). Some of the research suggests improved recall of information on summative tests after frequent exposure to information, even when formal feedback from a teacher was absent. When feedback was introduced, there were corresponding increases in outcomes (Kang, McDermott, & Roediger, 2007). The selection of frequency cycles in quiz administration should also be considered in the decision to use strands as a form of intervention. Frequent testing can be defined in various time cycles as follows:

- Long-cycle: across marking periods, quarters, semesters, years (length: 4 weeks to 1 year)
- Medium-cycle: within and between instructional units (length: 1–4 weeks)
- Short-cycle: within and between lessons (length: day-by-day; 24–48 hours; minute-by-minute; 5 seconds to 2; Wiliam & Thompson, 2007).

The testing effect can be combined with what is called the *spacing effect*. The spacing effect suggests that information is better received when offered over intervals of time as opposed to in one setting (Roediger & Karpicke, 2006). In support of the spacing effect, the predominance of research points to short- and medium-cycle formative assessment as having a positive impact on student learning to a greater degree than long-term benchmark or interim testing (National Council of Teachers of Mathematics [NCTM], 2007).

A number of studies have demonstrated the value of frequent formative assessment quiz administration (Bangert-Drowns, Kulik, & Kulik, 1991; Peat et al., 2004). For example, in a meta-analysis of 78 studies of low-, medium-, and high-interval testing, though significant differences in outcomes between intervals were not found, all outcomes were more beneficial to student achievement when compared to non-quiz-takers who participated in summative testing (Basol & Johanson, 2009). Another analysis included a study of assessment frequency in a course in civil engineering (14 quizzes) and a course in public health engineering (5 quizzes). The civil engineering course quiz frequency suggested a strong relationship between the number of quizzes taken and higher overall grades; in public health engineering, there was minimal suggestion of impact on learning (Aravinthan & Aravinthan, 2010). In another compilation of 29 studies, 13 of the studies showed a significant positive outcome association of having one quiz at 15 weeks. Only one of the studies reflected adverse results associated with the quiz addition (Bangert-Drowns et al., 1991). While studies show some promise in using quiz frequency to promote improved outcomes, there are also divergent opinions on the impact of such quizzes as assessment.

Opposition to the adoption of frequent quizzes in the formative assessment process echoes the concerns of overall formative assessment opponents—time taken away from curriculum learning is a very common complaint. Yet, the use of online formative assessment quizzes and the associated analysis tools may reduce such concerns as educators move from paper-based quizzes. Secondly, though formative assessment quizzes have low-stakes, some associate frequent test-taking, no matter what the size, with adverse impact on students (Marshall, 2007). Tests often have a poor connotation, whether formative or summative. Perhaps the distinction will become more widely known as formative assessment becomes more commonly used as a tool for improvements in student academics. Thirdly, opponents of frequent quiz administration assert that the process of testing in cycles does not lead to long-term retention but instead emphasizes teaching to the test only (Marshall, 2007). Though there are inconclusive research results on quiz taking for formative assessment (Gholami & Moghaddam, 2013), the research supporting gains in student achievement when using frequent quizzes provides optimism for public education use. Similarly, the flexibility of technology for classroom learning has contributed to alternative opportunities of instructional differentiation.

Computer-based instruction (CBI). Many public schools have turned to blended learning to meet the changing and growing need of educating a diverse population (Gulc, 2006). Although there are many methods of integrating technology into the classroom, blended learning is one means of combining the benefits of face-to-face (F2F) learning and online course content. For the purposes of this program evaluation, blended learning is defined as "a pedagogical approach that combines the effectiveness and socialization opportunities of the classroom with the technologically enhanced active learning possibility of the online environment, rather than a ratio of delivery modalities" (Dziuban, Hartman, Moskal, Sorg, & Truman, 2004, p. 3). Though there are many variations of blended learning, the teacher decides time spent in the traditional or online formats of computer-based application intervention. With the addition of blended learning, teachers can utilize online technology tools that automatically differentiate instruction, creating a customized learning experience for students. In the elementary schools of this study context, there are multiple computer-based and online mathematical learning tools made available to schools. Some software programs are mandated for use in all schools in the district, while others are selected and adopted by the principal of the school. In general, the use of computer-assisted learning provides a means of differentiation of instruction, encourages ongoing feedback, promotes student self-assessment, and motivates students to participate in the learning process (Dziuban et al., 2004; Gulc, 2006).

More than ever before, educational institutions are embracing a reality that teaching is not a "one size fits all" practice. Students have different learning styles, so it stands to reason that exposure to multiple ways of learning in the classroom may, in fact, improve academic results (Barbour & Reeves, 2009). In the study of context, the process of curriculum mapping is encouraged. This process identifies the teaching objectives and provides recommendation for the tools, to include technology tools, available to assist in meeting objectives in the classroom. Teachers turn to online technology application inclusion to reach students on multiple levels in the finite confines of class instructional time. One of the benefits of online applications is the availability of use outside of normal class time, extending learning exposure time for student. Though there have been studies that suggest much success in the integration of blended learning (Campbell, Gibson, Hall, Richards, & Callery, 2008), other studies suggest the difference in effectiveness between face-to-face and blended learning is negligible (Andrews & Haythornthwaite, 2007). Of the factors that impact the effectiveness of blended learning, there is some consensus in the following factors: (a) software selection, (b) frequency of feedback between teacher and student, and (c) administrative support (Watson, 2008). In incorporating blended learning, ready access to these elements should be considered when selecting this intervention.

The availability of technology-based formative assessment has grown. With the growth of technology-based assessments, advantages have surfaced that may combat some of the challenges faced in the formative assessment process. Teachers no longer have to develop tests in paper-and-pencil or on a computer, so readily available computer-based assessments reduce the time required to assess students. In addition, online formative assessments that are of best quality not only provide reporting on student results but provide feedback for steps toward intervention. With so many commercial assessments available that claim to be in alignment with state standards, the verification of such claims is generally left to the school division. Subsequently, the period required to verify standard alignment can be time intensive. The mathematics formative assessment in the state of Virginia, the ARDT, provides a unique opportunity for alignment with state summative assessments in that both the summative (SOL) and formative (ARDT) assessments for mathematics were developed by the same vendor and are offered to students in the very same graphic user interface (GUI). The convenience of an online formative assessment coupled with the assurance of summative test alignment make ARDT an attractive choice for a best practices tool in mathematics. Unfortunately, the need to address shortcomings in public schools associated with algebra readiness has a long history, but intervention options provide strategies for change (Brown, Hinze, & Pellegrino, 2008).

Tutoring programs. Tutoring has long been accepted as a means to provide students who show need an opportunity to receive assistance outside of traditional class time (McElvain & Caplan, 2001; Slavin, 1999). Unfortunately, securing tutors, either in-house or externally, may present a barrier to districts that are financially challenged. Title

1 grant funding is one of the ways the government supports the academic progress of disadvantaged and struggling schools, and these funds are heavily depended upon in the district of context to fund additional teachers as tutors. In schools that have been classified as Title 1 and are in 3 years or more of failure to meet adequate yearly progress (AYP), districts are required to offer parents access to supplemental educational services, which are more commonly identified as tutors (NCES, 2009). The selection of tutor-to-student model is a district prerogative. As such, there are many factors that are thought to contribute to success in tutoring (McElvain & Caplan, 2001; Slavin, 1999).

The effectiveness of tutoring programs has been attributed to the relationships that undergird them. Teachers who have existing and positive relationships with students are thought to be more influential to student growth (Hattie, 2009). District teachers as tutors not only have established relationships with students but are more familiar with the learning objectives of the curriculum and can more intently keep students fixed upon academic targets (Rothman & Henderson, 2011). Additionally, students exhibit more commitment to learning when teachers as tutors are perceived to be caring and familiar with the needs of the tutee (Cassellius, 2006; Klem & Connell, 2004; Triplett, 2004). In the VDOE-supported ARDT program of this study, one of the requested staff positions supplemented by VDOE funding is the classroom tutor. With district staff as tutors, teachers are positioned to encourage tutoring session attendance. Since attendance in tutoring programs is directly linked to academic gains (Hock, Pulvers, Deshler, & Schumaker, 2001; McComb & Scott-Little, 2003), having a tutor who can encourage attendance may contribute to higher participation in tutoring programs than those held after school. Yet, when utilizing existing teachers may not be feasible, there are other common characteristics that have been found to be important to the selection of tutors in general.

Tutors should possess academic and cultural aptitude in addition to the communication skills that have been found to contribute to successful progress in student tutees. To produce change in learning, master and certified teachers have produced higher impact in tutoring environments than volunteers or paraprofessionals (Gordon, 2009; Wasik, 1997). Having experienced educators as tutors who reside in or close to the school facilities where tutoring occurs is also thought to contribute to positive outcomes in tutoring sessions (Feldman & Ouimette, 2004). Whether volunteer or in-house teachers are acting as tutors, a common thread of importance linked to student performance is fostering continual communication between the classroom teacher and the tutor (Gordon, 2009; Gordon, Morgan, Ponticell, & O'Malley, 2004; Zuelke & Nelson, 2001). Additionally, other methods of tutoring such as peer tutoring show promise (Hattie, 2009).

Peer tutoring. Peer-assisted learning, also referenced as peer tutoring (Mastropieri et al., 2006), is a teaching approach that pairs students who have faced and conquered a learning task with those who are academically weaker or potentially face a challenge. This pairing approach to intervention has been associated with consistent and prompt gains in academic progress when compared to traditional teacher-centric settings (Hattie, 2009; Rohrbeck, Ginsberg-Block, Fantuzzo, & Miller, 2003; Roswal et al., 1995). Similarly, various organizations and researchers have noted the positive effect that peer-tutoring has had on mathematic aptitude in students (Fantuzzo, King, & Heller, 1992; Kunsch, Jitendra & Sood, 2007; National Tutoring Association, 2002). For example, one meta-analysis review of multiple forms of interventions for mathematics found overall peer assistance to be more effective than any other method studied (*d* = .60; Hartley, 1977). Two additional meta-analyses, one containing findings from 65 schools (Cohen, Kulik, & Kulik, 1982) and another evaluation of 19 studies (Cook, Scruggs, Mastropieri, & Casto, 1985) found significant effect sizes associated with peer tutoring and mathematics. Much of the success of peer assistance is attributed to a dependence on feedback from one student to another, which is faster and thought to be more impactful than traditional, unilateral teaching methods (Fuchs, Fuchs, Philips, Hamlet, & Karns, 1995). Depending on student need and resources available, peer assistance can be implemented in multiple forms.

Peer tutoring has been orchestrated and observed with students of like and differing ages, different subjects, as well as with students with disabilities. Peer teaching, when used by children with disability, was in many cases more impactful than the teaching methods that were routinely used (Mathes & Fuchs 1994; Telecsan, Slaton, & Stevens, 1999). Cross-age peer tutoring utilizes students who are older for tutoring younger students. The results of meta-analyses revealed that cross-age tutoring produced effect sizes of *d* = .79, slightly higher than the same-age peers of *d* = .59 (Hattie, 2009). Reciprocal peer tutoring is another form of intervention where the role of tutor and tutee are interchanged. One of the advantages of the reciprocal approach is the avoidance of superiority feeling in either the tutor or tutee, as well as the mutual benefits of shared roles. Further, an advantage for students is teaching with faster feedback, and a less threatening environment safe from whole-class judgement of student's learning styles or method of inquiry (Webb, 1988). As a result, both the tutor and tutee improve cognitive and academic skills (Galbraith & Winterbottom, 2011; Kalkowski, 2001). Likewise, tutors are thought to gain improved social skills, self-confidence, higher communication skills, and more in-depth subject matter knowledge (Mathur & Rutherford, 1991; NTA, 2002). Alternative peer tutoring uses class-wide peer tutoring such as Peer Assisted Learning Strategies (PALS) which are more formal, structured approaches that have foundation in whole class participation, awards for goals met, and deliberate teacher orchestrated pairing (Fuchs et al., 1995). All in all, the varied tutoring options offer multiple choices for districts that are supported by research. In the context of the study, the tutoring models that are most accessible to teachers on the local level are same-age and reciprocal tutoring. In recent years, the district has not supported a universal peer tutoring model, but such activities can be observed in schools with varied levels of application. In closing, the aptitude and effort of teachers to jointly identify and practice interventions meeting the needs of diverse populations may well hold the answer to how impactful formative assessment can be in the lives of students.

Summary

In summary, in a perfect educational setting, all students would academically progress towards higher mathematic aptitude, but in public, urban, low SES settings such results are not always a prevailing reality (Bailey et al., 2005; Vilorio, 2014). With many variables impacting student success, not only is early identification of student needs required, but the teachers' selection of how best to offer students interventions also becomes equally critical to the task of compounding academic progress. While technology-based formative assessment affords divisions and teachers ready access to assess student progress, the data alone is not the catalyst for change. The catalyst for change is the intervention selected by teachers that invokes a useful feedback stream from student-to-teacher, thereby accelerating the learning process (Hanover Research, 2016; Hattie, 2009).

CHAPTER 3

METHODS

Introduction

Formative assessment is a researched means of measuring student progress in the teaching process (Hattie, 2009). Yet, there remains a need to identify and apply evidence-based formative assessment interventions in districts serving diverse SES and ELL student populations (Alvarez et al., 2014; WIDA, 2009). In this program evaluation, I explored teacher intervention responses to ARDT data and the grade-level progress of students as indicated by the ARDT data. By understanding the overall impact of the ARDT formative assessment administration on student mathematic progress and the teachers' choices of intervention, districts may discover information to support improvement in the overall effectiveness of the ARDT program.

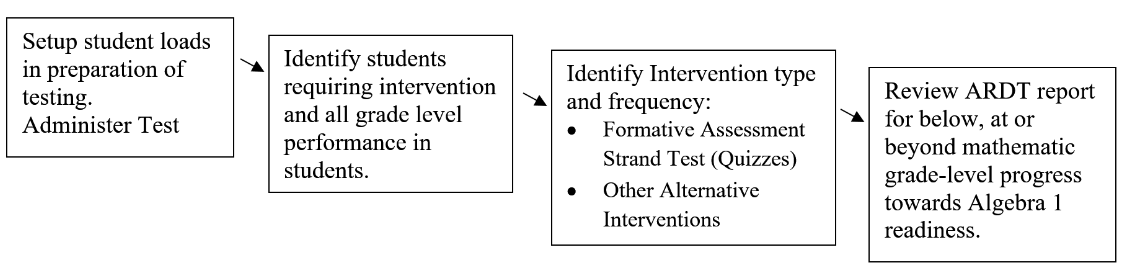

Figure 2. ARDT process model: Focus - intervention and grade-level progress.

This chapter outlines the methods of the program evaluation. First, the questions of the evaluation are described. Second, descriptions of the participants in the study are discussed. Next, the data sources and data collection methods used to support each question are explained. Thereafter, the data analysis choices for each question are detailed. The time line of this evaluation is also included in this chapter. Finally, the assumptions, limitations, and delimitations of the study are highlighted, and ethical considerations of the study are discussed.

The software vendor that provided the online mathematic formative assessment that were an integral part of ARDT program was Pearson. The formative assessment was adaptive, meaning the questions were altered in difficulty based on the student's response to questions. Additionally, the formative assessment contained technology-enhanced items (TEI) that allowed for multiple means of student interaction for answers to questions. When students completed an ARDT formative assessment, the system associated a 4-digit numerical score with the student's performance. This score was then cross-referenced to a grading table that aided in the classification of student performance as "at, below, or beyond grade-level." This classification of student performance was used to determine a student's grade-level algebra readiness.

There were three questions associated with this evaluation. Questions one and two involved discussions of grade-level algebra readiness, demonstrated by required testing and then grade-level outcomes from testing done above the required frequency, as reported in ARDT data. Question three involved the intervention choice used by teachers in response to data obtained from the ARDT reporting.

Questions. The evaluation questions I used for this program evaluation study were as follows:

RQ1: Is there a difference between fifth grade end-of-year mathematic scores start-of-year scores as measured by the ARDT?

RQ2: Does the frequency of ARDT formative assessment effect the algebra readiness of students as measured by ARDT reporting scores?

RQ3: How do teachers use the ARDT grade-level progress reports to guide instructional intervention choice?

Participants

Students. The elementary schools selected for this study were very similar in student demographic; however, one of the schools of the program did have a higher Hispanic population. All of the students in fifth grade participated in the ARDT process by taking the fall and spring assessment. This evaluation used the scores from all fifth-graders in each school. I selected fifth-graders because of the expected mathematical strength of students leaving elementary and entering middle school.

Knowing the mathematical strength of students entering middle school, administrators could compare the performance of feeder elementary schools. This data may be further explored to review interventions associated with success in schools that outperform others in student mathematics progress. The demographic makeup of school populations is included to increase the likelihood of comparisons in future use of the data produced in the study. Table 3 provides the student demographic data associated with each school of study.

Teachers and administrative staff. The inputs and talents of multiple members of the teaching and administrative staff were utilized. The educational experience and demographics of the teachers varied, but the majority of the mathematics teaching staff were female and held certification in their taught subject. Teachers new to the district during the study were provided training in ARDT over the summer. Throughout the school year, there were ARDT trainings held in schools for teachers to share new concepts and to reinforce knowledge transfer. These trainings, in general, were conducted by the Algebra Readiness Coordinator (ARC), who was the coordinator of the ARDT programs for all schools.

Table 3

School Demographic Makeup

School	Total students	Grade 5 students	Hispanic	Black	White	Asian	All Other
A	686	98	45.0%	51.8%	1.0%	0.7%	1.2%
B	442	71	29.0%	66.3%	2.7%	1.1%	0.9%
C	551	80	33.9%	59.9%	3.5%	1.5%	0.9%

Data Sources

Both quantitative and qualitative data were collected for this program evaluation. The formative assessment scores of student samples in three elementary schools were used to answer questions one and two with quantitative data. In addition, I conducted semi-structured interviews to generate qualitative data from teachers to answer question three. The questions and the associated data sources are found below.

RQ 1. Is there a difference between fifth grade end-of-year mathematics scores and start-of-year scores as measured by the ARDT? The formative assessment test is the heart of the ARDT program. This technology enhanced formative assessment enhanced assessment by making it readily available to teachers and offered immediate access to individual student results in the form of reports (see Appendix B). The first administered pre-assessment (fall) and post-assessment (spring) ARDT scores of fifth grade students from each school for 2016–2017 were collected to operationalize student progress from the opening to the ending of the school year. Both the pre- and post- ARDT tests consisted of 30 adaptive questions designed to measure student grade-level mathematics aptitude.

The ARDT formative assessment produced a 4-digit score for each student. The district then used a scoring map that assisted in the classification of each student in one of three categories based on a range of scores (i.e., at, below, or beyond grade-level expectation). The mapping was intended to communicate student progress in Algebra 1 readiness. Thus, the school ARDT fifth-grade student reports for both the beginning of the year and the end of the year were used as a data source.

RQ 2. Does the frequency of ARDT formative assessment administration effect the algebra readiness of students as measured by ARDT reporting scores? The ARDT captured the dates and frequency of each grade-level (30 question) assessment and content-strand assessment (10 questions) administered by location and by teacher. This ARDT assessment determined, at any point administered, if a student was assessed at, below, or beyond grade-level performance. The CAT grade-level test (30 question) was adaptive and was a larger-quantity representation of the content strands (10 questions) that covered five content areas: (a) numbers and number sense; (b) computation and estimation; (c) measurement and geometry; (d) probability and statistics; and (e) patterns, functions, and algebra.

Both grade-level and content-strand assessments could be used by teachers for intervention. The administration of one fall and one spring ARDT assessment was mandated by VDOE. All other ARDT assessment administrations, to include content strand testing and additional grade-level testing throughout the year, were discretionary and implemented as districts, principals, and teachers elected.

The ARDT report data associated with the total number of strand tests and grade-level tests combined, taken throughout the year for all fifth-graders from each school in the study, were analyzed for this study. Additionally, ARDT report data that provided all fifth grade final spring ARDT scores for each school of the study were used. This quantitative data was analyzed to determine if there was a significant difference in score progression between students who only participated in mandatory testing and those who participated in additional testing sessions. The data was reported as an overall comprehensive score representing each school.

RQ 3. How do teachers use the ARDT grade-level progress reports to guide instructional intervention choice? A qualitative interview script was used as the basis for a semi-structured interview. The interview questions were derived from the conceptual framework of the study, which emphasized formative assessment as a continual communication loop. The semi-structured interview contained questions that encourage the interviewee to freely express personal views and avoided interview questions that contain predefined answer choices. The interview script opened with a communication from the evaluator that explained the purpose of the study, introduced the evaluator's relationship to the context and study, and explained why the inputs of the interviewee are being sought. Also, informed consent notification was provided to participants and reviewed as a part of the interview in an effort to not only build trust but adhere to ethical research protocols. Further, the number of interview questions was held to a minimum (Creswell, 2014). There were five questions arranged in order from least intrusive to most difficult. Further, the interviewer asked additional probing questions to evoke more comprehensive discussions based on the answers offered to the primary five questions.

Probing questions supported emergent design in qualitative study in that such questions were designed to broaden conversations (Creswell, 2014). To interpret the information collected in the interviews, constant comparative analysis was used (Morgan, 1993). Each of the interviews was transcribed and reviewed to identify common ideas or themes. In each transcript of an interview, codes were assigned to various sentences, phrases, and paragraphs. Next, member checking was performed, in which interviewees reviewed responses and emergent themes derived from the process. The process of member checking increases the accuracy, credibility and validity of the findings in qualitative research (Creswell, 2014). The participants were contacted individually and given the opportunity to review the final analyses of the coding process from the interviews and to provide feedback on the accuracy. See Appendix C for qualitative interview questions.

Data Collection

To begin recruiting participants for the study, I contacted the ARDT lead mathematics instructional specialist of the district and conveyed the intent to conduct a program evaluation of ARDT. Next, I communicated the purpose of the program evaluation that is founded in identifying methods that improve student readiness for Algebra 1.

The ARDT coordinator had access to resulting data for all schools and served as the optimal provider of cross-location data produced in ARDT reporting and analysis. Finally, in a group setting, all teachers were invited to participate in the ARDT intervention interview. Those who agreed to participate were contacted for a scheduled one-on-one interview. Table 4 provides a summary and rationale for data that was collected.

Table 4

Data Elements and Collection

Question	Data Collection	Contribution
Is there a difference in fifth grade mathematics scores at the beginning of the school year and scores at the end of the year, as measured by the reported ARDT scores?	ARDT reports of resulting scores realized in fall (first assessment) and spring (final assessment) by school	Reporting data representative of student academic scores from beginning to end of school year, reflective of change in mathematics progress by school
Does the frequency of ARDT formative assessment affect the algebra readiness of students as measured by ARDT reporting scores?	ARDT reports of the total number of assessments administered in each school for the school year 2015–2016. Simple frequency data collection & ARDT reports of average final (spring) assessment scores by school	Correlation between test frequency and final scores earned in the spring (final) ARDT assessment for each school
How do teachers use ARDT grade-level progress reports to guide instructional intervention choices?	Semi-structured interviews conducted with teachers, tutors, and principals	Insight into intervention processes that key stakeholders identify as transformative

Data Analysis

The data collected in this program evaluation provided both quantitative and qualitative information. Mixed-methods research was considered because of its association with revealing information from stakeholders that may have not been considered prior and the inherent ability to validate findings using multiple tools (Creswell, 2014).

RQ 1. To run descriptive statistics for each school, I used a matched pairs t-test to determine if the differences between pre- and post-tests were statistically significant using an alpha of .05.

RQ 2. ARDT reports were reviewed, by school, revealing the total number of times the ARDT assessments were administered at each school during the school year of study. The student data was then separated into two groups: (a) students who participated in only two formative assessments, and (b) students who participated in more than two required formative assessments. Descriptive statistics were reported for each group. For all schools, I calculated a repeated measures ANOVA to determine if there was a statistical difference between scores between students who only participated in mandatory testing and those who participated in additional testing sessions. I selected the repeated measures ANOVA because RQ 2 introduced multiple characteristics over multiple observation points that might influence the outcome of the question. Additionally, a repeated measures ANOVA allowed for further analysis of data elements that might be considered relevant to RQ 2. For example, this analysis not only allowed for comparison of finished scores between the two groups, but also for comparison of starting mathematics aptitude of students for both groups.

RQ 3. Interview responses were analyzed for teacher opinions on intervention selection by conducting an initial analysis of the interview transcripts. Once completed, I performed member checking. Individual interview transcripts were sent to each interviewee, and they were encouraged to provide input for verification of initial data-gathering validity. Coding was conducted in order to identify any themes and topics that surface in the interview that are directly found in the literature review or problem statement. In this grouping of qualitative responses, similarities in interviewee opinions emerged, allowing for a more comprehensive reporting of findings. Table 5 provides a summary of data source and analysis with the questions these activities are intended to address.

Table 5

Data Sources and Analysis

Question	Data Source	Data Analysis
Is there a difference in fifth grade mathematics scores at the beginning of the school year and scores at the end of the year, as measured by the ARDT scores?	ARDT fall and spring test results for school year 2015–2016	Compare the comprehensive score from fall testing to spring testing for each school; matched pairs t-test will be used
Does the frequency of ARDT formative assessment administration affect algebra readiness of students, as measured by ARDT reporting scores?	ARDT reports of the total number of assessments administered in each school for the school year 2015–2016; overall comprehensive score growth for each school	Determine any correlation between test frequency and final scores earned in the spring (final) ARDT assessment for each school; use of repeated measures ANOVA to test for the equality of means
How do teachers use the ARDT grade-level reports to guide instructional intervention choices?	Semi-structured interviews conducted with teachers, tutors and principals	Qualitative analysis used coding as the primary source of data assessment

Table 6 depicts a timeline of the program evaluation tasks performed to produce the information gathered the evaluation process.

Table 6

Study Timeline

Timeframe	Activity
Late May	Retrieved spring and fall reporting from ARDT reporting system: Average student scores, total test number by school
Mid-June	Met with each principal; communicated purpose of study and sought approval to include school in the evaluation Met in group setting with mathematics teachers and tutors from each school; explained purpose of evaluation and requested individual time for one-on-one interviews
July	Conducted interviews
August	Data analysis and write-up

Assumptions and Limitations

Assumptions. Many education programs have their origin in perceived or known utility to address or avoid known problems. Assumptions in evaluations are those items that are a requirement for the research to be relevant. Changes of assumptions negate the need for research, because the absence of assumptions invalidates the problem, the primary driver for education research (Leedy & Ormrod, 2010). The first assumption for this evaluation was that Algebra 1 would remain the first secondary course that places emphasis on abstract mathematical reasoning, bridging primary mathematical and higher mathematical learning. The second assumption was that formative assessment would remain a highly predictive and research-supported tool for accurate measurement of student academic progress. The review of literature suggested that both assumptions were true: Algebra 1 was still the gatekeeper to higher mathematics (NMAP, 2008) and formative assessment was a highly recommended tool for detecting student progress for intervention (Hattie, 2009). Another assumption was that teachers, principals, tutors, and the ARDT coordinator will provide honest input in the interview process. To address this assumption, the interviewer explained the measures taken to maintain the confidentiality of information and protect the anonymity of the interviewee as outlined in the institutional review board (IRB) process.

Limitations. The samples in this evaluation were limited to three elementary schools in one school district and were a subset of the division's total elementary school population. The research supported the need to select locations that had similar demographic and socioeconomic status to normalize factors that might influence results that could influence the fidelity of implementation (Creswell, 2014). However, one of the schools did have a higher Hispanic population, which had bearing on the school's population of ELL students. Also, only students in fifth grade were included in this study. This population was an appropriate focus because the data associated with this population places emphasis on student mathematic readiness when transitioning to secondary school. This program evaluation assumed that measures of student grade-level achievement found in the ARDT system reporting were valid measures of student academic success. The VDOE comparison of ARDT and SOL test results showed consistency in their assessment results, providing another indicator to support the validity of the ARDT. Additionally, the teachers in some cases may have felt pressured to make claims that they are using multiple interventions, even though their intervention and differentiation selections may be limited. Preserving the confidentiality of the responses was meant to encourage participant trust. Lastly, the evaluator works in the technology department of the district of context, has over seven years' experience in education settings and is very project-, methodology-, and process-oriented. The bias associated with project management methodology and philosophy as a lens may have existed as a barrier to interpretation of processes performed in the ARDT program.

Delimitations were as follows:

- Context was in the state of Virginia;
- Context was a high-risk, public, urban elementary school;
- Total minority population exceeded 70%; and
- Readiness for algebra was determined by multiple measures, such as other formative assessments and SOL scores.

Ethical Considerations

The United States, on both federal and state levels, has persistently invested financial and human resources into programs that are intended to improve the educational experience for its K-12 students. These committed investments lend themselves to the use of a systemic means of evaluating return on investment (ROI) for monies allotted. The program evaluation standards produced in 1981, 1994, and 2011 by the Joint Committee on Standards for Educational Evaluation (JCSEE) defined a framework for evaluators that provides guidance in the form of standards that direct substantive program evaluation. According to the *Program Evaluation Standards: A Guide for Evaluators and Evaluation Users*, evaluation is a systematic investigation of the worth or merit of an object. The program evaluation standards are grouped into five areas: utility, feasibility, propriety, accuracy, and evaluation accountability. While there are five categories, each is understood to be of equal value (Yarbrough et al., 2011).

The categories of program evaluation standards provide guidelines for uniform methods for research. The area of utility seeks to support a program evaluation that produces information found credible and useful. Stakeholder input should be sought early and throughout the process to encourage involvement in the process. Utility speaks to the usefulness of results that are expected to be produced in the evaluation. Feasibility is meant to guide the efficiency and feasibility of the study, its human resources used and time expended. The evaluator conducted teacher surveys as a means to avoid potential impact to classroom time of instruction. The final three standards are propriety, accuracy, and evaluation accountability. The propriety standard governs ethical and legal responsibility in the program evaluation process; the accuracy standard guides data gathering and dissemination responsibility. These components govern the standards of program evaluation (Yarbrough et al., 2011).

Finally, the process for approval to conduct the study was twofold. The William & Mary Institutional Review Board provided input, direction, and initial approval of the study. Next, principals of the school were consulted for their personal approval to use their schools as a part of the context of the study. Lastly, the school division executive leadership conducted a review and committed final approval to conduct the study.

Summary

The chapter extends an overview of the methodology used to conduct the program evaluation of the Algebra Readiness Diagnostic Test (ARDT) in a Virginia school district. I collected data to support two of the three questions with quantitative information extracted from the ARDT reporting system, while the third question relied on qualitative measures to increase the overall validity of the evaluation and to discover valuable information in discussions with support personnel that have not yet been tapped. The purpose of this study has been to provide insight about the ARDT program's implementation and outcomes that will be useful to the practitioners in this context.

CHAPTER 4

RESULTS

The purpose of this study was to evaluate a technology-based formative assessment program intended to assist in intervention, remediation, and improvement of student outcomes in preparation for Algebra 1. The program selected was an online formative assessment, the Algebra Readiness Diagnostic Test (ARDT). In review, chapter 3 provided an overview of the methodology of the study, including the context, participants, data sources, and data analysis. Chapter 4 provides an overview of the study results and is organized by program evaluation question and relevant interview questions. Further, results of all quantitative and qualitative data collection gathered and associated with the study are described in this chapter. Below, the research questions guiding the evaluation are reviewed.

Evaluation Question 1 (EQ1): Is there a difference in fifth grade mathematic scores at the beginning of the school year and scores at the end of the year, as measured by the reported ARDT scores?

Evaluation Question 2 (EQ2): Does the frequency of ARDT formative assessment administration effect the algebra readiness of students as measured by ARDT reporting scores?

Evaluation Question 3 (EQ3): How do teachers use the ARDT grade-level progress reports to guide instructional intervention choices?

Paired T-Test

A paired samples t-test, supporting quantitative analysis, was used to answer Evaluation Question 1. The paired-samples t-test produced quantitative data that facilitated the comparison of student academic scores from the algebra readiness diagnostic test (ARDT) in the beginning of the school year to student academic scores of the ARDT obtained at the end of the school year. Student ARDT scores attained during the first assessment of the school year that began in September of 2016 and ended in June of 2017 were used. The number of students who participated in the beginning-of-the-year assessment and were also administered an end-of-the-year assessment are shown in Table 7.

Table 7

Number of Students Taking both Pre- and Post-tests

School Name	Student Count
School R	92
School F	82
School B	144

Repeated ANOVA

A two-way repeated measures ANOVA was used to answer Evaluation Question 2. Evaluation Question 2 sought to determine if there was a statistically significant difference between final ARDT scores of students who took more than three ARDT formative assessment tests and students who took three or less ARDT formative assessments. The number of ARDT assessments selected for comparison was chosen to be reflective of the number of required assessments given to students compared to the number of assessments given, by student, which exceeded required assessment administration.

Principal Interviews

One-on-one principal interviews were conducted on July 19, 2017, through July 20, 2017. Each of the three principals was an African American woman. Throughout this chapter, the respective principals will be referred to as Principal B, Principal F, and Principal R. The durations of principal interviews were as follows: Principal B., 18 minutes and 17 seconds; Principal F., 33 minutes and 6 seconds and Principal R., 25 minutes and 2 seconds. Interview questions were derived from the conceptual framework of formative assessment, which supports establishing where learners are in their learning and how to reach their learning goals **(**Wiliam & Thompson, 2007).

The principal interview was comprised of 8 questions. Question 5 of the 8-question interview was used to support Evaluation Question 2: Does the frequency of ARDT formative assessment administration effect the algebra readiness of students as measured by ARDT reporting scores? Questions 1, 2, 3, 4, 6, 7, and 8 of the principal interview were used to support Evaluation Question 3: How do teachers use the ARDT grade-level progress reports to guide instructional intervention choices? Specific principal interview protocols are below.

Principal Interview Questions

P1. Tell me about your experience or familiarization with formative assessment practices in mathematics.

P2. During the ARDT cycle, explain the processes that you use to determine the mathematics needs of our students in the fall of the school year.

P3. Explain how you use ARDT to identity intervention choice for students.

P4. Tell me the extent you utilize interventions in our classroom such as tutoring, peer-tutoring, CBI, assessment/test frequency or others?

P5. Explain your observations of student outcomes that occur as a result of assessment/test frequency.

P6. How do you identify the time you will spend in instruction verses the values of time spent in intervention?

P7. Tell me about any other forms of mathematics formative assessment quiz intervention that you use with frequency in the school year?

P8. Throughout the school, how do you communicate individual progress points to students?

Teacher Interviews

To provide insight into teacher initiated activities occurring in classrooms as a result of data obtained in ARDT reports, individual one-on-one teacher interviews were conducted at each school. There were seven teachers interviewed: 1 African-American man, 1 Caucasian woman, and 5 African-American women. Teacher interviews were conducted over two school days during the summer break, either after students had been dismissed or during student recess and lunch breaks so that instructional time would not be impacted.

The teacher interview pool consisted of four total questions: Question 4 was intended to address Evaluation Question 2. The remaining teacher interview questions, 1–3, were to answer Evaluation Question 3. Specific teacher interview protocols are below.

Teacher Interview Questions

T1. Throughout the school, how do you choose to communicate individual progress points in mathematics to students?

T2. Do you use student mathematic assessment reporting in your student planning decisions?

T3. Tell me about the extent you utilize interventions in your classroom such as tutoring, peer tutoring, Computer Based Instruction (CBI), assessment/test frequency or others?

T4. Is there a relationship between student mathematic assessment frequency and student mathematic learning?

The interviews with principals and teachers coupled with both statistical ANOVA and t-test helped to triangulate the data gathered to answer the questions of the program evaluation. The findings of this data are discussed in the pages that follow.

Evaluation Question 1: Is there a difference in fifth grade mathematic scores when compared between the beginning of the school year to the end of the year as measured by the reported ARDT scores?

Assessment results from 318 students spanning three schools were used to measure student performance in mathematics, and there was a significant difference, $t(310) = -11.20$, $p = 0.000$, between the scores at the beginning of the school year ($M = 1519.10$, $SD = 71.28$) and the scores at the end of the school year ($M = 1556.32$, $SD = 74.61$). Data suggest a difference between beginning-of-year and end-of-year fifth grade mathematic scores as measured by the reported ARDT scores (see Table 8).

Table 8

Mean Difference between Pre- and Post- Scores

		95% CI			
	M (SD)	Lower	Upper	t(310)	p
Prescore-Postscore	37.22 (58.60)	-43.76	-30.68	-11.20	0.000

Given that first- and end-of-year scores compare two points in time for student academic progress, a better qualitative understanding of the nature of additional assessment activities occurring between September and end-of-year would provide further insight into other impactful factors that may have been a contributing impetus to academic outcomes.

Evaluation Question 2: Does the frequency of ARDT formative assessment administration effect the algebra readiness of students as measured by ARDT reporting scores?

The results of a two-way repeated measures ANOVA revealed no relationship between frequent ARDT assessment and student outcomes. However, it should be noted that there were a total of 13 students from the three schools combined who participated in more than three ARDT test sessions. Though the ARDT mathematical content was aligned with the format and content of the mathematics Standards of Learning (SOL) summative assessment, very few teachers were assigning the smaller 10-question ARDT context strand test for reinforcement of skills, nor was the full ARDT administered with frequency outside of mandated assessments. Since there were very few students who participated in more than three assessments during the year, sample sizes for the ANOVA were much too small to establish a significant statistical relationship between overall frequency and mathematic progress in the participating schools. In contrast, the respondents of the interviews firmly believed that the time spent in frequent assessment increased student memorization of mathematical concepts to such a degree that the practice was thought to be vital to increased mathematical learning.

Principal interview responses. During separate interviews, each of the three principals described their building's formative assessment as a process that places heavy emphasis on time spent in planning building-level cumulative assessments after the mandatory fall ARDT assessment period. While the district's Pacing Chart provides a detailed analysis of expected days spent in the instruction of each mathematical concept, each principal expressed the need to also identify time in the classroom for re-visiting concepts taught in previous lessons throughout the year. For all the principals, for instance, cumulative tests in their schools consisted of concepts, in some quantity, taught from the first months in the school year to the month of the administration of the cumulative assessment. Since mathematical concepts are thought to be easier to obtain when previous mathematical concepts have been mastered, routine presentation of past concepts may serve as a catalyst to build stronger mathematical learning environments (Schiller et al., 2010). As such, the principals' responses to questions on the frequency of assessment reflected their unique instructional approaches.

Principal Question 5: Explain your observations of student outcomes that occur as a result of assessment frequency. As a result of the principals' communicated commitment to cumulative assessment frequency, their communicated support of classroom practices created a continual feedback loop between student and teacher that contributed to ongoing awareness of where students were in their knowledge of the subject matter for students, teachers, and principals alike. Principals believed that frequent assessment feedback gave them the opportunity to keep track of building progress, providing them a means to navigate to and sustain measurable academic progress. While each of the principals spoke to the need for assessments as a recall tool, two of the principals were more specific in their support for cumulative assessment use.

> Most of our assessments are cumulative, which we constantly review things that they should have learned in the past so they won't forget it. If you think about the math (as a subject), that builds each year so every year [we cumulatively assess]. If you do not constantly go over what they learned in previous years they are not going to make the connections. (Principal R)

To some degree, each principal equated much of their students' successes with the classroom teachers' practice of exposing students to problems or concepts cumulatively and doing so with intentional frequency; however, only one of the principals mentioned use of the ARDT strand test—but not in the traditional sense of administering strand tests to individual students. Instead, the principal commented that strand test questions provided classroom teachers a bank of problems that they utilized to develop and vary their own, site specific, cumulative assessments. In addition to the consensus that cumulative assessments should be at the foundation of schools mathematic teaching plans, each principal had also incorporated processes in the mathematics formative assessment process that encouraged sharing common content challenges and planning timely interventions.

None of the principals interviewed advocated ARDT strand or standard ARDT formative assessment as a preferred choice in our discussions of frequent assessment and academic outcomes. Instead, each principal spoke of the need to vary the ways in which students were assessed so that younger students would not grow disenchanted with the overuse of any one assessment approach. One principal described the need to find ways to engage younger students in ways that were viewed by students as fun, especially with elementary students. Variations in cumulative assessment mediums ranged from varying computer software such as Reflex, iReady, and MAP to a daily problem of the day or thumbs-down/thumbs-up activities in the classroom. In contrast, the ARDT assessment was not perceived as being engaging enough to maintain or garner the attention of elementary students, lacking animation and utilizing unchanging, stagnant content methods of standard delivery of elementary school assessment. In summary, all principals supported the use of mathematics assessments in varied platforms to increase the likelihood of keeping elementary students' vigilant during periods of assessment and to increase student achievement. But ARDT was not a preferred choice.

Teacher interview responses. Compared to principals' responses, the teachers expressed differing views on the value of frequent assessment during their interviews (see Table 9). Like the principals, the teachers conclusively felt that assessment frequency was a requirement for continual success in elementary mathematics. Without the frequency of assessment, teachers explained, children would likely forget content and falter in their progress.

Teacher Question 4: Is there a relationship between student mathematic assessment frequency and student mathematic learning? Similar to the principals interviewed, the frequency of assessment that the teachers endorsed also leaned heavily on the assessments that were developed and selected at their individual schools, not favoring sole ARDT use for mathematic assessment. Overall, the teachers interviewed expressed varied sentiments associated with the question of frequent assessment value.

A few of the teachers also mentioned using questions *from* the ARDT content strands in their routine assessment practices for group review, but also identified ARDT formative assessment content strand use as a customized response to individual student mathematic weaknesses in the classroom. One teacher had no doubt that using ARDT content strands was highly impactful. She explained that the professional development that she had received within the last two years had contributed to her confidence in using ARDT strands with individual students to help them with content mastery. This teacher had strong opinions that using ARDT content strands for students provided student confidence in the interface, which she believed translated into higher scores on the standardized summative assessment. She explained that, though she had only been in her current school for fewer than three years, she was not new to the district and that her excitement about ARDT was due to the impactful professional development she received at her current location.

> I have seen it with my eyes. [The students] have taken the ARDT. Their scores indicated an area that they are struggling in, we review, and I would say in two weeks they revisit. I assess them using the 10 questions from ARDT. The confidence level in them is like 10 times higher than when they first saw it. It's like then [the students] see it the second time it's like "Oh I know this." (Fifth Grade Teacher 2, Building R)

Common reservations. Along with the advantages of frequent formative assessments, both principals and teachers felt that there were inherent disadvantages associated with the use of frequent mathematic assessment. Some of the disadvantages mentioned in teacher interviews are paraphrased below.

Misleading assessment results. Due to very diverse populations of students who do not use English as their primary language, some students may know a problem but are prone to poor interpretation of unfamiliar words in a word problem. As a result, they may answer a problem wrong not because they did not know the concept, but because they were unfamiliar with the words associated with a number of assessment questions.

Impact of absenteeism. With overlapping assessments, multiple assessments occurring within one week, students who are absent may miss review of questions in multiple subjects. As a result, those children lack valuable classroom feedback and in turn risk falling behind faster. One teacher from School F explained that there was a belief that such continual failures damaged student confidence and that such confidence was hard to recover.

Multiple assessment burnout. Many of the required standardized mathematic assessments test the same content redundantly. There was concern that the need for multiple standardized assessments covering the same or similar content may not be completely justified and seemed to result in unnecessary student anxiety, burnout, and loss of focus.

Professional development alignment. Though ARDT has been supported by VDOE for many years, those years may have not always aligned with a district-wide investment in periodic yearly professional development, which would support teacher confidence in ARDT content strand test use.

In summary, Table 9 depicts common themes found in the teachers' opinions towards mathematic assessment frequency.

Table 9

Teacher Associations with Frequent Assessment

Themes	No. of Teachers
Harming Relationship (pressure, boredom, disengagement)	2
Correlation Relationship (identify strengths/weaknesses faster, early determination of student needs which benefits classroom teachers)	3
Review/Retention of Information (math/memorization connection)	3
ARDT Strand Preferred use for frequent mathematic assessment	1

An integral component of formative assessment success is the extent to which the data retrieved can be used to improve academic successes in students. As such, Question 3 of the program evaluation was intended to provide more intimate insight into activities that ensued as a result of ARDT report review. Again, both principals and teachers were interviewed with the primary objective of obtaining the opinion of people in both roles toward the ARDT formative assessment process.

How do teachers use the ARDT grade-level progress reports to guide instructional intervention choices?

Interview data gathered from principals and teachers revealed that both groups valued the practice of frequent assessment of students. Among all teachers and principals, there appeared to be an intentional effort to increase student exposure to math problems, but the methods that each selected were quite distinct. This continual process of seeking information through performance data and providing information among stakeholders in the learning process, student-to-teacher and teacher-to-student, mirrored the formative assessment conceptual framework supported by Wiliam and Thompson (2007), which also identifies the need for checkpoints.

Principal response to ARDT data. The reports retrieved from the online assessment of ARDT provided data on individual student mathematic performance in various content areas of mathematics. A teacher's pedagogical responses to such reports are thought to be highly influential to the mathematics outcomes realized by students (Hattie, 2009). These responses contribute to how teachers move to the established goal of student learning, the final component of the formative assessment conceptual framework, through the analysis of student formative assessment data.

Data meetings. There were three common themes that emerged in the responses of the principals. The first theme identified data meetings as a heavily valued practice used by each principal to stay apprised of academic progress in classrooms. These data meetings, for two of the principals (Principal R and Principal F), were a period for school staff to review student progress school-wide. These meetings, as expressed by the principals, were insightful of student progress but also served as an opportunity to review teacher performance. If the majority of students in a classroom were performing poorly in a content area, according to Principal B, teachers had a responsibility to find new ways to visit the lessons in re-teaching. Additionally, all of the principals reported that these data meetings allowed teachers an opportunity to exchange information on successful techniques used to deliver otherwise challenging content.

Daily assessment in instruction. A second theme that emerged was that the principals all viewed assessment as much more than a formal student evaluative performance. Principals spoke of their encouragement of daily assessment practices in classrooms. For example, Principals B and F referenced the use of daily problems, exit tickets, and differentiated homework assignments as regular classroom practice. All such practices were thought to be needed daily communications between students and teachers. Yet, the effects of the assessment process were not limited to the confines of the school building, and principals expressed appreciation for the impact of external factors on assessment success.

External factors. Third, all of the principals made mention of factors external to classrooms that had impact on student formative assessment results. Thus, a theme emerged in the data that assessment results are directly impacted by factors extending from the homes of students into schools.

Parental Involvement

Principal R talked about the importance of engaging parents in the process to promote student achievement, because elementary students seemed to perform better when they knew parents would be regularly informed of progress. Further, Principal R also felt that parent engagement was so important to student mathematical aptitude that her school made an investment in including parents in the assessment process. For example, School R sends home to families weekly newsletters with information on the classroom activities that occurred in the week and reports on the student's academic and behavioral performance. The newsletters are placed in envelopes made of materials that cannot easily be torn which have dates and lines on them for the parent's signature.

Additionally, School R has regular after-school meetings where parents are invited to have snacks and discuss student progress. According to all the teachers, home environmental factors in low-SES and high-ELL populations impact student performance in the classroom. Teachers mentioned the impact of newly incarcerated parents, high absenteeism, and threat of deportation on overall student performance. Principal F makes a commitment to pursue such factors especially when student performance drastically changes:

> There could be contributing factors; there are things that come out. Like… a student you see that fell into Tier 3, but he is usually a Tier 1 student. So then we ask the question, "Why is he Tier 3 with this assessment?" There could be different things. Like there was a situation once where the little boy was worried about deportation. This was a child that usually does well, but they noticed his demeanor change and once that fear was subsided he was back at Tier 1. (Principal F)

Thus, the principals considered the drastic change in a student's historical ARDT scores as possible contributing factors that could influence the formative assessment process.

Classroom responses. The ARDT assessment data can be a source to provide information on student mathematic needs but knowing weakness is only the start of attacking academic areas of limitation. Principals responded that they expected teachers to intervene or remediate based on assessment findings, but they did not view teachers as solely responsible for improving student performance. Each of the principals incorporated some degree of student responsibility for their own learning, promoting student ability, even at such young ages, to self-assess, make decisions about their paths to progress, and locate their motivation to excel academically. Thus, another common theme emerged that supported principal belief that students should play an active role in the formative assessment process.

Empowering students. Using the information gathered in ARDT, principal R's school shared the results of the assessments with students so students knew exactly how they did (based on percentage correct) and why (based on explanations of the answers given in class). Principal R had a large chart in a common area of the school that was divided into three colors: green, yellow, and red. Each child was assigned a number and allowed to periodically move the dot associated with their number depending on how they performed on the assessment. Green was the optimal color. Yellow suggested the child needed some improvement and a red classification communicated to the child that this was an area (subject) in which they needed to work much harder. Principal R explained that the children liked to promote their upgrade progress on the board and that the assignment of numbers reduced the possibility of any embarrassment for students. In short, these students knew where they were in the learning process and why at all times and made their own decisions on where more focus may be needed. Moreover, students were empowered to self-assess and navigate their own learning, which is also thought to be one of the more optimal ways to engage students in directing their possibilities for deeper learning (Klem & Connell, 2004). Overall, all principals used methods that invited students to actively participate in knowing the status of their own learning, but the methods used to reach students and encourage their participation did vary.

The one intervention that all principals discussed was the use of small groups. For instance, Principal B spoke of teachers creating small groups within the classroom based on the special needs of students that are identified in the results of ARDT reports. Conversely, Principal F referenced her schools use of the pull-out method that utilized temporary removal of students from their normal classroom setting to other classrooms that would address their needs in small groups or one-on-one tutoring sessions. Similarly, the push-in model was also mentioned where another teacher would come into the classroom with the classroom teacher and teach small groups in the primary setting.

During the interviews, principals referenced various technologies, peer tutoring, and frequent testing as methods used for interventions. Technology-based intervention answers surfaced in each principal answer. Again, multiple software tools were discussed as options for intervention. Imagine Learning, explained Principal R, was one of her favorite mathematic tools for ESL students because she felt that the students learned faster using the software program. Principal R stated, “My ESL children use Imagine Learning and we have iLead programs. So that works for the children. They put on the headphones; they are learning words, phrases. They catch on pretty quickly.”

The common theme among all principals was that the Pacing Chart dictated the baseline for how time should be spent in the classroom. However, they described various ways to align classroom time with the time periods for instruction in the Pacing Chart. With much content to cover and varied needs in large classrooms, the principals all had a regimen that they felt kept their schools on schedule. Principal F believed that instructional time was protected, but the work of intervention and remediation occurred simultaneously.

> Within the small group time they had with the students there was a teacher group, and usually that is the remediation group. So you may have four groups, and then

> there are three different rotations and so then Group 1 moves to Table A and then they will go to ABCD then they rotate around but there may be a small group of kids that the teacher needs to work with because they need intervention. This is how the teacher is both able to have a class and do the intervention at the same time. (Principal F)

Principal B, on the other hand, talked about having planning meetings with her teachers and mathematic specialists routinely to discuss student progress and share strategies that are working in content areas. Somewhat similarly, Principal R talked about the importance of intervention need driving mathematic instruction time due to the nature of mathematics concepts building upon each other year by year. In summary, all principals valued Pacing Chart schedules that provided guidance but the assessment data still drove how they would govern populations that needed special assistance concurrently.

Activities that ensued after assessment were many and varied by school, according to the principals interviewed. In feedback or in written documents, I could not identify guidelines that directed how students should receive information on their academic progress throughout the year outside of the report card. For each school, the way in which the principals decided to inform students was distinctly their own.

> At [School B] mastery is 90% so we are showing them a board that has, I think it starts at 50 and goes to 100. The students know exactly how they are doing on each assessment so they come down here with their assessment data and they have a fish this year. It was a horse in the past, but now they have fish and they move their fish towards mastery and they *want* to come down here with their teacher [to move the fish]. (Principal B)

Communicating student progress. Communications of student progress through visual cues seemed to be a popular practice with the three principals interviewed. Principal B and Principal R described areas in the building where all of the students came to post ongoing assessment progress. Periodically, students came to those areas in the building to move figures up or down on progress charts. Student anonymity was maintained for students by having the objects moved have their student number affixed upon them instead of names. Principal B spoke about the students' excitement to come down to the data wall and move their fish icon, representing assessment score progress. While all students may not have liked it, Principal B explained, this process provided an opportunity to speak to students that were underperforming and to inquire about what they could do differently to improve their scores for the next visit. In a similar manner, the teachers at School F, according to the principal, had data boards in each classroom. Another visual cue mentioned was the creation of human graphs. Principal R highlighted this classroom technique, used during assessment review, where students stood up if they answered a particular question correctly. Students were able to see how others faired on the same question and also, when correct, celebrate their successes. Additionally, students were encouraged to share problem solving techniques and receive the assistance they needed to master questions missed prior. Finally, Table 10 summarizes the interview responses of principals.

Table 10

Reported Practices by Principal

Practice	Principal B	Principal F	Principal R
Experience and familiarization with assessment	Daily classroom assessment practices (exit tickets, problem of the day)	Bi-weekly data meetings Daily classroom assessment practices (exit tickets, problem of the day External contributing factors to student assessment performance	Bi-weekly data meetings External contributing factors to student assessment performance
Fall preparation activities for math instruction	ARDT assessment to group and tier students Review of 4th grade assessments for 5th grade students	ARDT assessment to group and tier students	Review of 4th grade assessments for 5th grade students Professional development to build teacher confidence in recognizing student academic needs
Interventions selected based on ARDT Results	Did not use ARDT reports to determine interventions (alt. prog. used: Interactive Achievement, ARDT strand questions)	Small group composition and tier classification Multiple software learning applications (Tail Gate Mathematics, Sunrise Mathematics Academy, Mouth Full of Mathematics) Visual aids depicting each student's ARDT progress (anonymized)	Small group composition and tier classification
Classroom interventions	Small group (pull outs) Bi-weekly cumulative assessments Computer-based software (Reflex)	Peer tutoring (pair share, turn and talk) Small group (pull outs) Computer-based software (Reflex, Kool Math)	Small group (pull outs, pull ins) Computer-based software (Imagine Learning, Brain Pop)
Ways students monitor progress	Data walls/data boards	Human graphs Personal binders	Data walls/data boards
Intervention vs. instruction time	Instruction time driven by pacing chart	Instruction time driven by pacing chart	Interventions drive instruction time

Teacher Response to ARDT Reports

Throughout the school, how do you choose to communicate individual progress points in mathematics to students? The teachers used ARDT report data in some of the same applications identified by principals but also discussed their unique means of data use. Not surprisingly, one of the more prevalent ways teachers communicated individual ARDT progress to students was through the sharing of assessment results. Two of the teachers expressed a special affinity to finding time to have one-on-one sessions for review of assessment data with students in their classrooms, no matter what the class size. Consequently, according to the teachers who favored student grouping, the assessment results facilitated their decisions to group students according to aptitude demonstrated on the ARDT. Many assessment software programs not only provided individual assessment results but also classroom averages. When this information was shared with students, Teacher F2 believed that a comradery ensued where students worked collectively to improve the classroom pass-rate. According to him, students helped each other more and wanted to push each other to do better when they knew that everyone was actively involved in the classroom average. Additionally, many of the traditional ways of communicating to students were also ways that were identified to communicate to students.

Others expressed commitment to communicating student academic progress in varied ways. For instance, one teacher, R3, associated communicating progress to students with increasing student ability to set goals, which she believed to be a requirement for their deeper learning. Similarly, another teacher, F2, wanted to make certain that her students knew that her standard for their learning was not 70% but 80%;

her expectation was that students would also learn to set goals for baselines higher than the minimum standard of 70% required by the Standards of Learning, Virginia's statewide accreditation assessment. "The data tell us how much time we really need to spend on specific skills and how much time we don't have to spend" (Teacher F1). Table 11 summarizes teacher use of ARDT data for communicating.

Table 11

Communicating Student Progress with ARDT Data

Communication activity	Teachers reporting activity
Technology Assessment Feedback Classification	F3, R1, F2, B1, B2
Setting Goals High	B2, F2
Small groups/ 1 on 1	F3, R1

Do you use student mathematic assessment reporting in your student planning decisions? For some teachers, using assessment reports began early in the school year as they sought to identify student strengths and weaknesses. Teacher B1 explained that all of the fifth-grade teachers in the school received the fourth-grade mathematics assessment scores of their students. Once the reports were received, Teacher B1 further explained, decisions were made to select the most appropriate class content to assist in any weaknesses found in the fourth-grade report. Teacher C2, similarly, spoke about the selection of ARDT strand content choice being driven by assessment data. Similarly, review of ARDT reporting provided pass percentages on the entire classroom, which gave the teacher an opportunity to self-critique and determine if re-teaching was necessary.

Student groupings. Teachers also used assessment report data to drive how their classroom groupings would occur. One of the teachers (Teacher F2) explained how large class sizes, with upwards of 30 students, required grouping students to deliver effective teaching. Students were grouped so that the stronger students could act as leads with weaker students, a form of peer tutoring. In elementary schools, Teacher F2 explained, this was a good time to begin to introduce leadership skills to students. Moreover, some of the teachers opted to use both small group and peer tutoring in classrooms that were driven by assessment data. In summary, assessment data results directed time spent in mathematical concepts and consequently how students that needed assistance would receive it. Principals also enable these classroom decisions by providing the organizational support for the needed adjustments teachers much make (see Table 12).

Table 12

Mathematic Assessment Data and Teacher Planning

Influence on planning	Teachers reporting
Determines the time spent in mathematical content areas	At, B1, B2, F1
Review and retest	A3
Small group composition	A2, B2

Tell me about the extent you utilize interventions in your classroom such as tutoring, peer tutoring, Computer Based Instruction (CBI), assessment/test frequency or others?

Peer tutoring. Teachers had a strong sense of the need to use ARDT mathematic reports to assist them in guiding the instructional needs of students. Throughout the interviews there were shared aims expressed by the teachers in meeting the needs of ethnic, age, and size diverse populations.

Even as teachers discussed other choices used for intervention during their interviews, by far, peer learning was highly favored among teachers. In fact, one teacher's (A1) only response for intervention choice was peer tutoring. On multiple occasions, teachers expressed the need to be creative in identifying ways to make certain that every child received what was needed in large student-to-teacher ratio settings. For some of those teachers, peer tutoring offered another way for teachers to include differentiation in their classrooms through student teachers. When the teacher's methods of instruction were not well-received by a student, another student may be able to explain the topic in a way that learning would still occur: "I love to partner a strong student with a weaker student. Sometimes I feel they may not get it from me but they might get it from another kid" (Teacher B2).

> What happens is one child is reviewing that skill and helping the other kid to get that skill so it's strengthening the peer as well as the student with that weakness. We have a large population of Hispanic kids here so lots of times we will get kids that speak very little English and so we will have to match them with kids that know Spanish to help them understand and interpret. (Teacher F1)

Peer tutoring seemed to be an intervention choice that all the teachers endorsed not only to benefit the academically weaker student, but also to begin to develop leadership skills in the stronger student. Likewise, another choice common for intervention among teachers was computer-based instruction (CBI).

Computer Based Instruction. The teachers in this program evaluation used many different kinds of CBI (see Table 13). For example, one teacher (B2) spoke about using ARDT strand test content in the classroom within rotating small groups throughout the class to reinforce learning by content strand. Additionally, Reflex Mathematics was a software program that each school, to some degree, had endorsed as a program for both interventions and remediation efforts. Teachers mentioned the flexibility in reporting that the software contained which was a characteristic that added to the draw of students' seemingly enjoying their time in the application. One teacher (C2) explained that CBI was reserved for 20 minutes of each class day, so CBI time was incorporated into the lesson plan. Equally important, there were other themes that emerged in answers to this interview question.

Table 13

Teacher Intervention Selections

Intervention	Teachers reporting preference
Peer Tutoring	B2, C1, C2,F1
Computer Based Instruction	
Reflex Math	B2, F3
iReady	B2
First in Math	B2
Prodigy	F3
Slumdog	F3
Small Group	F3, B2,B1

Tutoring. Teachers also have opinions about intervention choices that they did not feel were impactful. One of the teachers spoke briefly on her opinions of tutoring (Teacher C1). She felt that tutoring was a useful tool in helping students but expressed apprehension about the effectiveness of volunteer tutors. Additionally, she felt that student academic improvements were greater when the selected tutors had relationships with the students and were more familiar with the student's needs. In the same manner, Teacher C1 had concern for time lost in the classroom associated with pull-outs. According to Teacher C1, when students were not in the class during pull-outs, students that were already behind were missing valuable class time. There were efforts being made to make up the time students were not in the classroom but there was a concern that pull-out students had a very hard time with the time lost during regular class time. One of teachers (B1), serving as a Title I specialist, expressed that she did not utilize peer tutoring because, overall, the students in that class were weaker academically. The preferred choices of intervention were peer tutoring and CBI.

Summary of Findings

In summary, the teachers of this Virginia district used mathematic reporting data from the ARDT formative assessment to govern their school days, identify student strengths and weaknesses, and continually assess students. The ARDT data from the previous year, when available for students, provided teachers a viable means to know student needs at the very beginning of the school year, allowing teachers to plan interventions that catered to the special needs of students all year long. In classrooms where student numbers were high, this information saved teachers time and was another means to decrease the likelihood that a student's needs would not be identified early due to lack of viable data.

Using Algebra Readiness Diagnostic tests, teachers were incorporating information into classroom lesson plans in such a way that minimal time was lost where intervention was needed. Additionally, as computer-based instruction (CBI) allowed teachers to differentiate instruction, knowing student ARDT performance contributed to teachers' ability to select software that would more likely help students become stronger academically. Similarly, CBI provided teachers a means to engage young minds longer, reducing the threat of boredom and behavioral mishaps in classrooms. Finally, the ARDT strand test questions, though not used in totality but only for the content of the questions in cumulative assessments, was also being used to differentiate instruction and target students in one-on-one efforts. The participants in this study found ARDT to be a useful tool in monitoring student academic progress in mathematics. As a complement to other instructional strategies, ARDT was beneficial for teachers who made the early identification of student needs as a priority in promoting higher probability of increasing academic successes and ultimately stronger skill demonstration in Algebra 1.

CHAPTER 5

DISCUSSION

The purpose of this program evaluation was to review the standardized mathematic formative assessment, Algebra Readiness Diagnostic Test (ARDT), to determine how teachers and principals are using the tool and data reports to prepare students for success in Algebra 1. In this chapter, I will provide recommendations to inform practitioners that are framed in the findings of the evaluation, associated with the problem presented in this program evaluation, and founded in the literature review. The implications of the study will be discussed, and recommendations for further research will be considered. Finally, I will offer my summary of the Algebra Readiness Diagnostic Formative Assessment program evaluation effort.

Recommendations

The journey in developing elementary students who will possess mathematic strength begins with purposeful steps towards providing appropriate resources and encouraging belief systems that support mathematic aptitude. The recommendations offered are motivated by a goal of informing practitioners of those practices that offer prospect to adequately prepare elementary students for the higher-level math courses ahead.

Adopt a Program Evaluation Process for Formative Assessments

A process for periodic evaluation of formative assessment software should be adopted in the division. Low-SES districts are often saddled with the responsibility of acquiring funding to meet numerous needs that are associated with diverse populations.

Without question, formative assessment software is often an expensive investment for school divisions. Such purchases obligate districts to yearly costs such as software licensing, training, and ongoing technical support. These investments are made with the expectancy of increasing positive outcomes in student achievement. In districts such as the one that is the subject of this program evaluation, there are multiple formative assessments that are being used for mathematics and English assessment and perhaps with good reason. Teachers expressed that students become disengaged with predictable content; however, there should be checks and balances in the form of program evaluation in place to support justification for continued investment in selected mathematics formative assessment software. The program evaluation would monitor teacher/administrator feedback and statistical outcomes associated with the use of the software. Doing so would serve to eliminate or reduce the possibility of routine yearly investments in a formative assessment that could be an ineffective platform or conversely support continued investment in a formative assessment that contributes to positive and measurable outcomes.

Select Age Appropriate Elementary School Mathematic Assessments

Elementary students are still developing their ability to stay focused in classrooms. As a result, teachers continually seek to identify engaging practices that will keep students actively involved and progressively learning. One of the observations of this program evaluation was that students were not as focused during the ARDT assessments as when using some of the other, more animated or content varying, online assessment software. In comparison, the ARDT provides a more sterile environment using primarily traditional black font for mathematics problems viewed on a white background that rarely changes. Online assessments for elementary students should be more age-appropriate, favoring content that is offered with graphics that are more inviting and less predictable for young audiences. In the selection of assessments and learning tools for elementary students, the user-friendliness of the selected software should promote elementary student engagement by providing an inviting learning experience catering to the very students who will use the software most. Otherwise, the impact of student boredom and disengagement may be a prevailing factor in the scores and as such might potentially impede a teacher's ability to make effectual decisions to support students.

Prioritize Feedback to Students

The selection of a quality mathematics formative assessment tool is essential to the success of schools. Successful instruction in schools is directly anchored to the practice of contiguous feedback (Hattie, 2009). While formative assessments inform teachers and districts of student academic levels, feedback from formative assessments provides students with another learning opportunity to address problem-solving techniques, encourage self-assessment, and support goal setting. Without reviewing the results of an assessment soon after it is taken, students may miss an opportunity to correct computations made in error and to subsequently learn from those errors. There are many software vendors that offer assessments; some offer feedback to students after the assessment. The investment in software that provides mathematic assessments to districts is not often a small one. Software vendors can be challenged to provide not only large banks of questions for student practice but also time sensitive student feedback paths in the form of explanations and answers to questions after the assessment. Given the frequency of standard assessments, students are missing a valuable opportunity for additional quality learning time when assessment answers are not reviewed. The time for participation in assessment is mandated; time-sensitive periods of missed question review should also be added as an integral part of the formative assessment process. Moreover, use of assessments that do not include answer share periods for students should be kept to a minimum or, optimally, eliminated. After all, the results of formative assessments are only as good as the quality of the assessment, the teacher's commitment to use the feedback offered toward modification, and the student's ability to receive and respond to reported progress.

Use Data to Adjust Instructional Priorities

During interviews, teachers communicated that they knew of certain areas in mathematics that were consistently harder for some students than others. One area mentioned as a problem area for students was word problems. In order to have data to substantiate the assertion that students are performing poorly in some areas, the district would want to identify the types of questions that have been consistently missed in previous years by analyzing past assessments. Once these problem areas are identified, mathematics specialists and teachers could identify a bank of questions that addresses these areas. Additionally, *there should be an intentional effort to identify and teach words that students classified as English Language Learners do not know during mathematic assessments*. Then, the principal would assign a period during the day or week when such problems should be presented for review of methods or steps to resolve such problems. For instance, after recess, there might be a 15- to 20-minute period where each teacher would be required to teach on the problem areas that have been noted with the bank of approved questions that have been created to build student strength in that area. This should be a time when teachers are practicing explicit teaching by demonstrating step-by-step problem resolution, which is a recommended practice for improving student mathematical skill (Hanover Research, 2016).

An additional area of struggle for elementary school students is learning their multiplication tables. During conversations with the ARDT coordinator of this evaluation, he explained, as evidenced in ARDT scores, how many students were entering and some leaving middle school still weak in multiplication aptitude. In short, students in inner city schools may reach eighth grade without having mastered their multiplication tables. This fact alone will impact their chances of going further in mathematics and succeeding at passing Algebra 1. Of course, reaching middle school and not knowing their multiplication tables, such students are likely embarrassed and shy away from participating in classroom mathematics activities. In response, there have been successful methods of introducing and reintroducing multiplication tables for mastery in a way where children learn through recall and are not embarrassed in the learning process. Districts should seek to incorporate cost conscious and proven methods in the classroom early in elementary schools to develop strong foundations in multiplication. For instance, Thomas Caron (2007) uses a simple sheet that provides students the answers to multiplication tables at the top of the paper with common multiplication tables below. Students are timed each week while answering these questions. The results of this practice have shown that students are more readily memorizing multiplication and are able to build upon this solid foundation in developing mathematic confidence.

Conduct Frequent Cumulative Assessments

Without review of mathematical concepts, student could potentially forget mathematic concepts that build upon each other. End-of-year mathematic Standards of Learning (SOL) content and ARDT formative assessment are geared to compile all required knowledge areas to assess student overall aptitude. Consequently, in order for elementary school students to prepare to be mathematically stronger, there must be periods of instruction where assessments contain samples of information taught to date. For example, Principal B described her school's assessment as *growing longer and longer each time,* which were the terms used to explain the increase in assessment question number that ensued as the year progressed.

Cumulative assessments should be conducted on the schedule that principals identify for their buildings, because principals are more likely to know what curriculum time adjustments must be made. All of the principals expressed their commitment in leading compliance to the mandated testing calendar for assessments provided to the district. To a smaller degree, each principal also communicated slight challenges associated with doing so. Though the Curriculum Pacing Chart for the district is intended for districts to cover content at the same time, the needs of students are very different, and some may require re-teaching on certain subjects, while others may not. Said differently, each classroom of students will have different needs, where some might take longer or shorter to learn content areas than another classroom. Mandated assessments that come between the first and last assessments of the school year assume that certain content has been taught prior to the assessment. In some cases, if a classroom is a few days behind, the mandated assessment may cover content that the class has not learned or reviewed. In short, teachers should have some autonomy to adjust classroom time based on the students' needs.

In order to accomplish this, districts could initiate an action research effort whereby principals would monitor the progress of classrooms where teachers have been encouraged to develop their own teaching calendars with principal oversight. These calendars would include dates of cumulative review, assessment, and feedback opportunities for students. The results of these assessments would also provide teachers an opportunity to provide differentiated homework to students to support strengthening concept aptitude in students as needed.

Cultivate Elementary Teachers' Mathematics Aptitude

Recruiting and retaining quality teachers for at-risk schools can be challenging (Stronge, 2010). Thus, teachers who are recruited may not possess strong mathematical skill. Therefore, elementary teachers are not likely to do well at teaching a subject in which they are not well-versed. In turn, students exposed to mathematics teachers who may be unable to demonstrate elementary mathematics skills will also transition to secondary schools with the limited depth of mathematics taught them.

Teacher competency in elementary mathematics should be continually cultivated since much of student future success in mathematics may depend on developing strong foundational understanding of mathematics in elementary school. One way to do this might be to require that elementary teachers, who do not have mathematics certifications or degrees, participate in a mathematic assessment or refresher training for the grade-level that they teach each year. The mathematic assessment used should be an online assessment that offers answers, explanation, and review of missed problems. Having teachers assess, districts would know the readiness of teachers to impart strong foundational knowledge into students.

The talents of teachers who have strong mathematics backgrounds should be used differently. If a school has been unable to acquire sufficient elementary teachers possessing strong mathematical backgrounds, the strong mathematics teacher should be shared on a schedule in all classrooms to safeguard that each classroom is receiving adequate time exposure to support mathematic foundations in students. As a result, students may benefit from more classrooms where teachers are providing explicit and heuristic-driven instruction (Jayanthi et al., 2008). In summary, mathematics refresher courses would keep teachers abreast of the knowledge that is expected to be conveyed in classrooms. To hope or project that students do well in mathematics classes leading to algebra and not require teachers to demonstrate such aptitude is to hope that students are exposed to and learn something that their teachers may not know themselves.

Provide Inducements for Teaching in At-risk Schools

Some research results suggest that at-risk students perform better when reared in homes where the caregiver is a college graduate (Davis-Kean, 2005). Unfortunately, high-risk students are not often raised in homes where the parent is a college graduate and are unlikely to be provided adequate mathematical assistance at home. Thus, having students be taught in schools with well-equipped teachers, in some cases, may be the only opportunity for a student to fully grasp mathematic concepts. Since elementary teachers teach multiple subjects, and since mathematics students may be more inclined to pursue science, technology, engineering, and mathematic (STEM) careers that do not include teaching, divisions must be creative in pursuing mathematically astute teachers to recruit for elementary schools. For instance, to attract quality teachers to districts that do not have adequate resources, tuition reduction or grants may encourage quality teachers to make a reciprocal investment in at-risk districts while providing graduates needed jobs in education. There are likely very few graduates who would not find the possibility of erasing school loan debt as an incentive to teach, even if only temporarily.

To pursue preparing college students as potential teachers for low-SES districts, colleges and K-12 districts will want to closely collaborate to produce teaching talent that will be most beneficial for the needs of students and produce teachers who are equipped to teach in diverse settings. As such, there should be vertical alignment between the course requirements for elementary school students and the course requirements for college students participating in the program. The college students enrolling in the program should be made aware of the special opportunity that the curriculum will provide them to promote elementary students' mathematical knowledge in such a way that those strengths will propel elementary students to stronger mathematic aptitude demonstration in secondary school. In short, in order for teachers to respond to the data found in mathematics formative assessment reports, teachers must be able to differentiate mathematics instruction in such a way that struggling students demonstrate improved understanding. To do so, teachers better serve the needs of students when they have strong mathematic aptitude.

Offer Ongoing Professional Development for Formative Assessment Data Use

Teachers should participate in mandatory bi-yearly professional development in mathematic formative assessment so that they are intimately familiar with the features that may benefit them in the classroom. A quality mathematic formative assessment

Table 14

Summary of Findings and Recommendations by Evaluation Question

Question	Findings	Recommendations
EQ1: Is there a difference in fifth grade mathematic scores when compared between the beginning of the school year to the end of the year as measured by the reported ARDT scores?	There is a statistically different score when comparing start of school to end of year ARDT scores	Research any additional contributing factors that might have impacted the ARDT score difference observed.
EQ2: Does the frequency of ARDT formative assessment administration effect the algebra readiness of students as measured by ARDT reporting scores?	Frequent cumulative assessment schedules in schools are thought to help students retain information.	Allow mathematic problem failures that are common to students to influence time spent in instructional assessments for years that follow.
		Cumulative mathematic assessments should be considered an integral part of local practice and conducted on schedules recommended by the principal.
EQ3: How do teachers use the ARDT grade-level progress reports to guide instructional intervention choices?	Teachers and Principals have differing opinions on the fit for use of the ARDT assessment.	Adopt a school or division-wide program evaluation process for mathematic assessment software.
	The ARDT school processes for student feedback differs and may have impact on outcomes for students.	Only invest in online formative assessments that contain feedback opportunities for students.
	There may be common mathematical challenges among elementary school students.	Use Mathematic assessment report data to adjust instructional content for the classroom.

allows a reporting mechanism for identification of missed questions and makes recommendation of questions for students to strengthen identified areas. A few teachers in this study were aware of formative assessment context strand tests but there was some indication that extracting content strands to differentiate instruction for individual students was not common practice for them. In contrast, one of the teachers who seemed very comfortable with strand test administration commented that in the last two years she has become familiar with the tool through training with the Algebra Readiness Coordinator. She also expressed her high satisfaction with the benefits she had realized using them to make students stronger academically. As another cost-saving measure, the Instructional Technology Resource Teachers (ITRT) could be trained by the ARDT coordinator or mathematics instructor in a train-the-trainer model. All associated costs for the professional develop effort could potentially be associated with existing staff time. The ITRT staff, who are in schools throughout the school week, could utilize their talents to provide refresher courses for teachers and hone teachers' skill-level in using mathematic formative assessment features.

Implications for Leadership

School districts should support planned forums that encourage principals with proven success models to share with other principals of the same grade levels. Principals whose schools have a history of reaching mathematic accreditation in a larger diverse population should participate in bi-yearly sharing sessions with other principals. In school B, the principal commented on internal sharing of strategies among teachers as being a strength in the school. There were processes that all of the schools were using to orchestrate school-wide activities in mathematical learning. Much of the information shared at monthly meetings with principals can be shared in a newsletter, leaving more time for meaningful professional development about teaching and learning. Principal meetings, as an alternative, can be strategizing sessions while district news is shared in other forms of communication. It might be that some schools are performing at a slightly lower level but a few changes in practice might make the difference in students' mathematics achievement outcomes.

To make change, principals would lead breakout sessions for their grade level, review their internal process, and compare/contrast processes with other school principals. For principals whose schools were struggling, that principal should be encouraged to incorporate variations or new practice in his or her school routine. At the meeting's end, the principals would report to executive leadership the modifications or additions that they are planned. The executive leadership, in turn, would communicate a timeline for such activities to be implemented and request interval reporting from the principal. Slight or more profound changes may make the difference in failing mathematic scores for some schools and principals should be given an opportunity to learn from each other in a non-threatening and supportive forum.

Implications for Education

In order for students to enter secondary school with a strong foundation in mathematics, districts must be more strategic in preparing students in elementary school. The challenging role played by elementary teachers in leading students to mastery in multiple subjects to include mathematics cannot be understated. To support teachers in providing the best academic leadership for students, school districts should consider investment in mathematics and assessment professional develop for teachers that are versed in mathematics as well as those that have lower mathematic self-efficacy. Additionally, diverse student populations benefit academically when multiple ways to learn are offered to them (Gulc, 2006). In the program evaluation findings, for instance, teachers explained the reason for using so many technology assessments in the classroom was that students easily become bored and disengaged when the same technology was placed before them. As such, teachers varied the methods used to reach students academically by using varied teaching methods and tools.

In order for formative assessments to deliver the value that is expected by districts, prospective programs must be carefully scrutinized to make certain that such assessments are aligned with the appropriate standards, that the problem banks offered in the software tool are several, that the formative assessment tool will provide feedback on next steps to assist students, and that reporting mechanisms are user friendly. Additionally, content should be evaluated by the subgroup of teachers and principals who will be using them. Since principals and teachers are closest to the learning process, their input on the formative assessments selected should be a critical criterion in its selection.

Discovered Limitations of the Study

Beyond the required testing periods of ARDT, there were very few students who were administered the context strand assessments. Thus, the samples were too limited to conclusively address the impact of frequent testing as intended to answer question EQ2 (*Does the frequency of ARDT formative assessment administration effect the algebra readiness of students as measured by ARDT reporting scores?*) In addition, a larger teacher sample would provide a wider-lensed feedback opportunity toward teacher use of formative assessment data in their classrooms. Further, this study does not associate teacher interventions in the classroom with individual teacher classroom mathematics results.

Recommendations for Future Research

Further research should be conducted on the impact of teacher use of heuristic mathematics practices on student formative assessment scores. Another research effort would be to investigate the comparison of weekly online formative mathematic assessment to monthly or bi-monthly formative mathematic assessment outcomes. This study could include multiple vendor formative assessment tracking as seems to be the more common practice of use in the schools where this program evaluation was conducted. Finally, another recommendation for research would be to investigate weekly parental communications for student mathematics progress, as compared with minimal feedback prior to standard report card dissemination.

Conclusion

A global society that is highly dependent on workers who have performed well in K-12 STEM courses will be unforgiving of students who are ill-prepared for work in an economy fueled by technological advances. Performing well in mathematics is far from optional for students; mathematic aptitude is a necessity for many quality post-secondary career opportunities (Domina, 2014). Mathematic formative assessment, if widely utilized, is a viable tool for monitoring student progress and designing instructional responses. By increasing mathematic concept recall, cumulative assessments in elementary may also be beneficial to student retention in later grades. Yet, formative assessment and teacher efforts alone cannot bridge the student mathematics achievement gap (Duncan & Magnusan, 2011; Stone, 1998; Walston & McCarroll, 2010).

The path for success in mathematics for students not only includes formative assessment feedback to teachers but also continual feedback to students and parents, who can partner with schools in the interventions selected to address the individual needs of students. We cannot underestimate the value of allowing students to self-assess. Students are able to self-assess when the results of formative assessments are made available to them, when they are encouraged to set goals throughout the formative assessment process, and when they are given opportunities to relearn concepts that were problematic for them in prior assessments.

It is important for educational leaders to understand that, for best results, the formative assessment process should extend beyond the assessment to a communication and action plan involving students and families. If given the proper support systems, elementary students can transition to secondary school highly equipped for mathematic learning leading to Algebra 1 and beyond. The formative assessment tool used can continue to provide the vital feedback needed for the primary stakeholders in student learning to make the needed adjustments to forge a pathway to more manageable navigation in student mathematic acceleration. However, the report should now be one of many opportunities for parents to be brought into the learning process for their students. Online formative assessments allow for easy access to reporting data that can and should be shared with parents and students with frequency. Closing the loop of assessment feedback to include elementary students and their parents may mean the difference between a student without a solid foundation and a student who has the encouragement, support, and expectation to achieve mathematic aptitude that will set them up for success in Algebra 1.

Appendix A

Informed Consent

Informed Consent to Participate in a Research Study

Title of Study: A Program Evaluation of a Technology Based Formative Assessment

Study Sponsor: The College of William & Mary

Investigators:

Name:	Dr. Margaret Constantino	**Dept:**	Education	Meconstantino@wm.edu
Name:	Dr. Michael DiPaola	**Dept:**	Education	mfdipa@wm.edu
Name:	Dr. Jacob Joseph	**Dept:**	Education	jdjoseph@wm.edu
Name:	Ms. Cassandra Harris	**Student:**	Graduate Program	charris@rvaschools.net

Introduction

- You are being asked to be in a research study: A Program Evaluation of a Technology Based Formative Assessment for Algebra Readiness.
- You were selected as a possible participant because of your participation in schools where formative assessment programs are executed.
- We ask that you read this form and ask any questions that you may have before agreeing to be in the study.

Purpose of Study

- The purpose of the study is to evaluate a technology-based formative assessment program intended to assist in intervention, remediation and improvement of student outcomes in preparation of Algebra 1.
- This research will be presented in partial fulfillment of the requirement for the degree of Doctor of Education.

Description of the Study Procedures

- If you agree to be in this study, you will be asked to do the following things:
  (1.)Participate in an interview.
  (2.)Provide honest and open feedback on all questions.

Risks/Discomforts of Being in this Study

- The study has no known risk

Benefits of Being in the Study

- The benefits of participation are the possibility of identifying algebra readiness process components that are beneficial to preparing students for strong competency in Algebra 1 and mathematic courses that follow Algebra 1.

Confidentiality

- This study is anonymous. We will not be collecting or retaining any information about your identity.

- The records of this study will be kept strictly confidential. Research records will be kept in an electronic file. The information will be coded and secured using a password protected file. At the completion of the student, the code will be electronically destroyed. We will not include any information in any report we may publish that would make it possible to identify you.

Payments

- There is no direct form of compensation being offered for participation in this study.

Right to Refuse or Withdraw

- The decision to participate in this study is entirely up to you. You may refuse to take part in the study *at any time* without affecting your relationship with the investigators or the College of William & Mary. Your decision will not result in any loss or benefits to which you are otherwise entitled. You have the right not to answer any single question, as well as to withdraw completely from the interview at any point during the process; additionally, you have the right to request that the interviewer not use any of your interview material.

Right to Ask Questions and Report Concerns

- You have the right to ask questions about this research study and to have those questions answered by me before, during or after the research. If you have any further questions about the study, at any time feel free to contact me, Cassandra Harris at cdharris@email.wm.edu or by telephone at 804-873-3051. If you like, a summary of the results of the study will be sent to you. If you have any other concerns about your rights as a research participant that have not been answered by the investigators, you may contact Dr. Margaret Constantino, Director of Executive Programs at meconstantino@wm.edu.

Consent

- Your signature below indicates that you have decided to volunteer as a research participant for this study, and that you have read and understood the information provided above. You will be given a signed and dated copy of this form to keep, along with any other printed materials deemed necessary by the study investigators.

Subject's Name (print): ______________________________

Subject's Signature: ______________________________ Date: ________________

Investigator's Signature: ______________________________ Date: ________________

THIS PROJECT WAS FOUND TO COMPLY WITH APPROPRIATE ETHICAL STANDARDS AND WAS EXEMPTED FROM THE NEED FOR FORMAL REVIEW BY THE COLLEGE OF WILLIAM AND MARY PROTECTION OF HUMAN SUBJECTS COMMITTEE (Phone 757-221-3966) ON 2017-06-19 AND EXPIRES ON 2018-06-19.

Appendix B

Algebra Readiness Diagnostic Test

Algebra Readiness Diagnostic Test (ARDT)
Tests, Reports, and Score Interpretation Information for Teachers

Revised 4.2.2013

Test and Report Information

Grade-level Tests – This is the first diagnostic test administered to each student. These 30-item, computer-adaptive tests assess content mastery for a particular grade level. Computer adaptive tests are tests delivered on a computer where the set of test items administered to students will vary based on how they respond to each item during the test. As a result, the 30-item Grade-level Test may include questions from previous grade levels if the student is not demonstrating mastery at the current grade-level.

Grade-level Test results may assist in determining the need for student intervention and in identifying specific content strands where instructional remediation would be beneficial. Two types of score reports are available for the Grade-level Tests:

- **On Demand Group List Report**
  - Provides individual performance overall and by strand for each student in a group
- **On Demand Student Detail Report (for teachers and parents)**
  - Provides bar graphs of performance overall and by strand for a student
  - Provides data for use in determining need for intervention/remediation
  - Provides specific item information including the SOL assessed by the item, an estimate of item difficulty, and whether the student answered correctly

Strand Tests – Strand Tests can be assigned to a student after he/she has taken the Grade-level Test. The computer-adaptive Strand Tests are 10 items in length and focus on a particular content strand. Strand Test results can be used to determine a student's strengths and areas of challenge within the content strand. Two types of score reports are available for the Strand Tests:

- **On Demand Group List Report**
  - Provides individual performance for a specific grade-level content strand for each student in a group
  - Provides specific item information including the SOL assessed by the item, an estimate of item difficulty, and whether the student answered correctly or incorrectly
  - Provides links for the teacher to view individual items from the strand test
- **On Demand Student Detail Report (for teachers and parents)**
  - Provides individual student performance information found in the Group List Report

General Scoring Information

Grade-level Test Scoring

Grade-level Test Administered	Content included	Possible Grade-level Test Scores	"On Track" Score	
			Beginning of School Year	End of School Year
CAT6	Grades 3-6	1300-1699	1550	1650+
CAT7	Grades 3-7	1300-1799	1650	1750+
CAT8	Grades 3-8	1300-1899	1750	1850+
CATAlg	Algebra I	1900-1999	N/A	1950+

Important General Notes

- The time of year that the ARDT is administered will have a direct impact on the interpretation of ARDT scores.
- A 6th-grade student that scored near 1550 at the beginning of the year is "on track" for the study of 6th-grade content.
- A 7th-grade student that scored near 1650 at the end of grade 6 or at the beginning of grade 7 is "on track" for the study of 7th-grade content.

Sample Grade-level Test Score Interpretation Scenarios

EXAMPLE 1

Sample Grade-level Test Group List Report (6th Grade) – Given on February 15, 2013

Student Name	STI	# attempts	Pre-test Score	Post-test Score	Best Score		
XXXXXXXXX	STI	1	1467	1467	1467	Number & Number Sense	1497
						Computation & Estimation	1396
						Measurement & Geometry	1430
						Probability & Statistics	1520
						Patterns, Functions, and Algebra	1477

Students studying 6th-grade content who are "on-track" for success should have performance levels that vary based on the time of the year that the ARDT was given.

- Beginning of the year – scores near 1550
- Middle of the year – scores between 1550 and 1650, depending upon exposure to content
- End of the year – scores near 1650

Based solely upon this ARDT assessment, possible observations might include:

- The student's scores in each content strand are below grade level. Since the student scored below 1550 (expected at the beginning of the year), he/she may benefit from intervention and remediation.
- Compared to the other content strands, the student's highest scoring strand (but still below grade level) is Probability and Statistics. A question to consider is whether the content assessed in this strand had been taught at this point in the year (February).
- The student's greatest weaknesses are in the Computation and Estimation and Measurement and Geometry strands. Intervention and remediation may target these areas first.

EXAMPLE 2

Sample Grade-level Test (7th Grade) – Given on February 15, 2013

Student Name	STI	# attempts	Pre-test Score	Post-test Score	Best Score		
XXXXXXXXX	STI	1	1705	1705	1705	Number & Number Sense	1742
						Computation & Estimation	1696
						Measurement & Geometry	1698
						Probability & Statistics	1665
						Patterns, Functions, and Algebra	1719

Students studying 7th-grade content should have performance levels that vary based on the time of the year that the ARDT was given.

- Beginning of the year – scores near 1650
- Middle of the year – Middle of the year – scores between 1650 and 1750, depending upon exposure to content
- End of the year – scores near 1750

Based solely upon this ARDT assessment, possible observations might include:

- The student's content strand scores are each near 1700, indicating that the student may be "on track" and may not need intervention based upon this assessment alone.
- In each content strand, this student seems to be "on track," with the exception of Probability and Statistics. A question to consider is whether the content assessed in this strand had been taught at this point in the year (February).
- Depending upon whether instruction in Probability and Statistics has occurred, the teacher may or may not determine that intervention is necessary.

Appendix C

Interview Protocols

Interview Questions: Principal

1. Tell me about your experience or familiarization with formative assessment practices in mathematics?

2. During the ARDT cycle, explain the processes that you use to determine the mathematic needs of your students in the fall of the school year?

3. Explain how you use use ARDT reports to identify intervention choices for students?

4. Tell me about the extent you utilize interventions in your classroom such as tutoring, peer-tutoring, CBI, assessment/test frequency or others?

5. Explain your observations of student outcomes that occur as a result of assessment/test frequency?

6. How do you identify the time you spend in instruction verses the value of time spent in interventions?

7. Tell me about any other forms of mathematics formative assessment quiz intervention that you use with frequency in the school year?

8. Throughout the school, how do you communicate individual progress points to students?

Interview Questions: Teachers

1. Throughout the school, how do you communicate individual progress points to students?

2. Do you use student mathematic assessment reporting in your student planning decisions?

3. Tell me about the extent you utilize interventions in your classroom such as tutoring, peer-tutoring, CBI, assessment/test frequency or others?

4. Is there a relationship between student mathematic assessment frequency and student mathematic learning?

1. Throughout the school, how do you communicate individual progress points to students?

2. Do you use student mathematic assessment reporting in your student planning decisions?

3. Tell me about the extent you utilize interventions in your classroom such as tutoring, peer-tutoring, CBI, assessment/test frequency or others?

4. Is there a relationship between student mathematic assessment frequency and student mathematic learning?

References

Abrams, L. M. (2007). Implications of high-stakes testing for the use of formative classroom assessment. In J. H. McMillan (Ed.), *Formative classroom assessment: Theory into practice* (pp. 79–98). New York, NY: Teachers College Press.

Alvarez, L., Ananda, S., Walqui, A., Sato, E., & Rabinowitz, S. (2014). *Focusing formative assessment on the needs of English language learners*. San Francisco, CA: WestEd.

American College Testing. (2015). *The condition of college & career readiness—African American students 2015.* Retrieved from www.act.org/research

Andrew, M. (2014). The scarring effects of primary-grade retention? A study of cumulative advantage in the educational career. *Social Forces, 93*(2), 653–685.

Andrews, R., & Haythornthwaite, C. (2007). *The Sage handbook of e-learning research*. Los Angeles, CA: Sage.

Antoniou, P., & James, M. (2014). Exploring formative assessment in primary school classrooms: Developing a framework of actions and strategies. *Educational Assessment, Evaluation, and Accountability*, *26*(2), 153–176.

Aravinthan, V., & Aravinthan, T. (2010). Effectiveness of self-assessment quizzes as a learning tool. *Proceedings of Engineering Education Conference*, Birmingham, UK, July 6–8.

Bailey, T., Jenkins, D., & Leinbach, T. (2005). *Community college low-income and minority student completion study: Descriptive statistics from the 1992 high school cohort*. New York, NY: Teachers College, Columbia.

Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C. L. C. (1991). Effects of frequent classroom testing. *The Journal of Educational Research*, *85*(2), 89–99.

Barbour, M. K., & Reeves, T.C. (2009). The reality of virtual schools: A review of the literature. *Computers and Education, 52*(2), 402–416.

Başol, G., & Johanson, G. (2009). Effectiveness of frequent testing over achievement: A meta-analysis study. *Journal of Human Sciences*, *6*(2), 99–121.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation, and Accountability 21*(1), 5–31.

Brown, J., Hinze, S., & Pellegrino, J. W. (2008). Technology and formative assessment. *21st Century Education*, *2*, 245–255.

Bryk, A. S., Lee, V., & Holland, P. B. (1993). *Catholic schools and the common good*. Cambridge, MA: Harvard University Press.

Camara, W. (2013). Defining and measuring college and career readiness: A validation framework. *Educational Measurement: Issues and Practice*, *32*(4), 16–27.

Campbell, M., Gibson, W., Hall, A., Richards, D., & Callery, P. (2008) Online vs. face-to-face discussion in a Web-based research methods course for postgraduate nursing students: A quasi-experimental study. *International Journal of Nursing Studies 45*(5), 750–59.

Caron, T. (2007). Learning multiplication: The easy way. *Clearing House: A Journal of Educational Strategies, Issues, and Ideas, 80*(60), 278–282.

Carraher, D. W., & Schliemann, A. D. (2007). Early algebra and algebraic reasoning. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (pp. 669–706). Charlotte, NC: Information Age.

Cassellius, B. (2006). Using relationships, responsibility, and respect to get from “good to great” in Memphis middle schools. *Middle School Journal*, *37*(5), 4–15.

Cates, G. L. (2005). Effects of peer versus computer-assisted drill on mathematics response rates. *Psychology in the Schools*, *42*(6), 637–646.

Clark, I. (2011). Formative assessment: Policy, perspectives, and practice. *Florida Journal of Educational Administration & Policy*, *4*(2), 158–180.

Cohen, P. A, Kulik J. A., & Kulik, C., (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal, 19*(2), 237–248.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Department of Education.

College Board. (2013). *Total group profile report: 2013 college-bound seniors.* New York, NY: Author.

Complete College America. (2016). *Corequisite remediation: Spanning the completion divide.* Retrieved from http://completecollege.org/tag/minority-serving-institutions

Cook, S. B., Scruggs, T. E., Mastropieri, M. A., & Casto, G. C. (1985). Handicapped students as tutors. *The Journal of Special Education*, *19*(4), 483–492.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, *66*(3), 227–268.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Cross, C. T., Woods, T. A., & Schweingruber, H. (Eds.). (2009). *Mathematics learning in early childhood: Paths toward excellence and equity*. Washington, DC: The National Academies Press.

Curry, K. A., Mwavita, M., Holter, A., & Harris, E. (2016). Getting assessment right at the classroom level: Using formative assessment for decision making. *Educational Assessment, Evaluation, and Accountability*, *28*(1), 89–104.

Darling-Hammond, L. (2010). *Evaluating teacher effectiveness: How teacher performance assessments can measure and improve teaching*. Washington, DC: Center for American Progress.

Darling-Hammond, L. (2015). *The flat world and education: How America's commitment to equity will determine our future*. New York, NY: Teachers College Press.

Darling-Hammond, L., & Sykes, G. (2003). Wanted, a national teacher supply policy for education: The right way to meet the "highly qualified teacher" challenge. *Education Policy Analysis Archives*, *11*(33).

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology, 19*(2), 294–304.

DeAngelis, K. J., White, B. R., & Presley, J. B. (2010). The changing distribution of teacher qualifications across schools: A statewide perspective post-NCLB. *Education Policy Analysis Archives*, *18*(28).

Domina, T. (2014). The link *between* middle school mathematics course placement and achievement. *Child Development*, *85*(5), 1948–1964.

Duncan, G. J., & Magnuson, K. (2011). The nature and impact of early achievement skills, attention skills, and behavior problems. In G. J. Duncan & R. Murnane (Eds.), *Whither opportunity? Rising inequality, schools and children's life chances* (pp. 47–69). New York, NY: Russell Sage Foundation.

Duncombe, C., & Cassidy, M. (2016). *Weighing support for Virginia's students: Virginia lags in providing low-income students the opportunities to succeed in the classroom*. Richmond, VA: The Commonwealth Institute.

Durwood, C., Krone, E., & Mazzeo, C. (2010). *Are two algebra classes better than one? The effects of double-dose instruction in Chicago* [Policy brief]. Chicago, IL: Consortium on Chicago School Research.

Dustin, D. S. (1971). Some effects of exam frequency. *The Psychological Record*, *21*, 409.

Dziuban, C., Hartman, J., Moskal, P., Sorg, S., & Truman, B. (2004). Three ALN modalities: An institutional perspective. In J. Bourne & J. C. Moore (Eds.), *Elements of quality online education: Into the mainstream* (pp. 127–148). Needham, MA: Sloan Center for Online Education.

Epper, R., & Baker, E. (2009). *Technology solutions for developmental mathematics: An overview of current and emerging practices* [Report]. Denver, CO: William & Flora Hewlett Foundation, Bill & Melinda Gates Foundation.

Fantuzzo, J. W., King, J. A., & Heller, L. R. (1992). Effects of reciprocal peer tutoring on mathematics and school adjustment: A component analysis. *Journal of Educational Psychology*, *84*(3), 331.

Feldman, J., & Ouimette, M. (2004). *Examining the Turning Points comprehensive middle school reform model: The role of local context and innovation*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA. Retrieved from https://eric.ed.gov/?id=ED509779

Flores, A. (2007). Examining disparities in mathematics education: Achievement gap or opportunity gap? *The High School Journal*, *91*(1), 29–42.

Foster, J. E. (1993). Retaining children in grade. *Childhood Education*, 70(1), 38–43.

Fuchs, L., Fuchs, D., Phillips, N., Hamlett, C., & Karns, K. (1995). Acquisition and transfer effects of classwide peer-assisted learning strategies in mathematics for students with varying learning histories. *School Psychology Review*, *24*(4), 604–620.

Gaertner, M. N., Kim, J., DesJardins, S. L., & McClarty, K. L. (2014). Preparing students for college and careers: The causal role of Algebra II. *Research in Higher Education*, *55*(2), 143–165.

Galbraith, J., & Winterbottom, M. (2011). Peer tutoring: What's in it for the tutor? *Educational Studies*, *37*(3), 321–332.

Gates, A. I. (1922). *Recitation as a factor in memorizing* (No. 40). New York, NY: Science Press.

Gholami, V., & Moghaddam, M. M. (2013). The effect of weekly quizzes on students' final achievement score. *International Journal of Modern Education and Computer Science*, *5*(1), 36.

Gordon, E. E. (2009). 5 ways to improve tutoring programs. *Phi Delta Kappan*, *90*(6), 440–445.

Gordon, E. E., Morgan, R. R., Ponticell, J. A., & O'Malley, C. J. (2004). Tutoring solutions for No Child Left Behind: Research, practice, and policy implications. *NASSP Bulletin*, *88*(638), 59–68.

Gulc, E. (2006). *Using blended learning to accommodate different learning styles*. Retrieved from Higher Education Academy website: http://escalate.ac.uk/downloads/2917.pdf

Hanover Research. (2016). *Best practices in at-risk secondary math instruction.* Retrieved from http://gssaweb.org/wp-content/uploads/2016/06/Hanover-Research-Spotlight-Best-Practices-in-At-Risk-Secondary-Math-Instruction.pdf

Hartley, S. S. (1977). *Meta-analysis of the effects of individually paced instruction in mathematics*. Boulder, CO: University of Colorado.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112.

Heppen, J. B., Walters, K., Clements, M., Faria, A. M., Tobey, C., Sorensen, N., & Culp, K. (2011). *Access to Algebra I: The effects of online mathematics for Grade 8 students* (NCEE 2012-4021). Washington, DC: National Center for Education Evaluation and Regional Assistance.

Herlihy, C. M., & Quint, J. (2006, November). *Emerging evidence on improving high school student achievement and graduation rates: The effects of four popular improvement programs* [Issue brief]. Washington, DC: National High School Center.

Hinton, V., Stroizer, S., & Flores, M. (2015). A case study in using explicit instruction to teach young children counting skills. *Investigations in Mathematics Learning,* *8*(2), 37–54.

Hock, M. F., Pulvers, K. A., Deshler, D. D., & Schumaker, J. B. (2001). The effects of an after-school tutoring program on the academic performance of at-risk students and students with LD. *Remedial and Special Education*, *22*(3), 172–186.

Jacobs, V. R., Franke, M. L., Carpenter, T. P., Levi, L., & Battey, D. (2007). Professional development focused on children's algebraic reasoning in elementary school. *Journal for Research in Mathematics Education*, *38*(3), 258–288.

Jayanthi, M., Gersten, R., & Baker, S. (2008). *Mathematics instruction for students with learning disabilities or difficulty learning mathematics: A guide for teachers*. Portsmouth, NH: RMC Research Corporation Center on Instruction.

Johnson, B. C., & Kiviniemi, M. T. (2009). The effect of online chapter quizzes on exam performance in an undergraduate social psychology course. *Teaching of Psychology*, *36*(1), 33–37.

Kalkowski, P. (2001). *Peer and cross-age tutoring* (School Improvement Research Series). Retrieved from Education Northwest website: http://www.nwrel.org/scpd/sirs/9/c018.html

Kang, S. H., McDermott, K. B., & Roediger, H. L., III. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology*, *19*(4-5), 528–558.

Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of School Health*, *74*(7), 262–273.

Kunsch, C. A., Jitendra, A. K., & Sood, S. (2007). The effects of peer-mediated instruction in mathematics for students with learning problems: A research synthesis. *Learning Disabilities Research & Practice*, *22*(1), 1–12.

Langdon, D., McKittrick, G., Beede, D., Khan, B., & Doms, M. (2011, July). *STEM: Good jobs now and for the future* (ESA Issue Brief #03-11). Washington, DC: U.S. Department of Commerce.

Lee, J., Grigg, W., & Dion, G. (2007). *The nation's report card: Mathematics 2007– National Assessment of Educational Progress at Grades 4 and 8* (NCES 2007-494). Washington, DC: National Center for Education Statistics.

Leedy, P., & Ormrod, J. (2010). *Practical research planning and design* (9th ed.). Boston, MA: Pearson.

Linacre, J. M. (2000). Computer-adaptive testing: A methodology whose time has come. In S. Chae, U. Kang, E. Jeon, & J. M. Linacre (Eds.), *Development of computerised middle school achievement tests* (MESA Research Memorandum 69). Seoul, South Korea: Komesa.

Martinez, J. G. R., & Martinez, N. C. (1992). Re-examining repeated testing and teacher effects in a remedial mathematics course. *British Journal of Educational Psychology, 62*, 356–363.

Marshall, B. (2007). A crisis for efficacy? *Education Review, 20*(1), 29–35.

Marshall, B., & Drummond, M. J. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education*, *21*(2), 133–149.

Mastropieri, M. A., Scruggs, T. E., Norland, J. J., Berkeley, S., McDuffie, K., Tornquist, E. H., & Connors, N. (2006). Differentiated curriculum enhancement in inclusive middle school science: Effects on classroom and high-stakes tests. *The Journal of*

Special Education, *40*(3), 130–137.

Mathes, P. G., & Fuchs, L. S. (1994). The efficacy of peer tutoring in reading for students with mild disabilities: A best-evidence synthesis. *School Psychology Review, 23*(1), 59–80.

Mathur, S. R., & Rutherford, R. B., Jr. (1991). Peer-mediated interventions promoting social skills of children and youth with behavioral disorders. *Education and Treatment of Children*, *14*, 227–242.

McComb, E. M., & Scott-Little, C. (2003). *After-school programs: Evaluations and outcomes*. Greensboro, NC: SERVE.

McCombs, J. S., Augustine, C. H., Schwartz, H. L., Bodilly, S. J., McInnis, B., Lichter, D. S., & Cross, A. B. (2011). *Making summer count: How summer programs can boost children's learning.* Santa Monica, CA: RAND.

McElvain, C. K., & Caplan, J. C. (2001). Creating effective after-school programs for middle and high school students. *NASSP Bulletin*, *85*(626), 35–44.

Mertens, D. M., & Wilson, A. T. (2012). *Program evaluation theory and practice: A comprehensive guide*. New York, NY: Guilford Press.

Morgan, D. (1993). Qualitative content analysis: A guide to paths not taken. *Qualitative Health Research, 3*(1), 112-121.

National Assessment of Educational Progress. (2015). *The nation's report card: 2015 mathematics and reading assessments.* Retrieved from https://www.nationsreportcard.gov/reading_math_2015/files/2015_Results_Appendix_Math.pdf

National Center for Education Statistics. (2009, November). *Statistics in brief: Students'* *use of tutoring services by adequate yearly progress status of school* [Issue brief]. Washington, DC: Author.

National Center on Educational Outcomes. (2012, July). *Including students with disabilities in common non-summative assessments* (NCEO Brief No. 6). Minneapolis, MN: University of Minnesota.

National Council of Teachers of Mathematics. (2007). *Five key strategies for effective formative assessment* [Research brief]. Retrieved from https://www.nctm.org/uploadedFiles/Research_and_Advocacy/research_brief_and_clips/Research_brief_04_-_Five_Key%20Strategies.pdf

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel* [Report]. Washington, DC: United States Department of Education.

National Tutoring Association. (2002). Peer tutoring factsheet. Retrieved from http://www.ntatutor.com/peer-tutoring-facts.html

No Child Left Behind Act, Pub. L. No. 107-110. (2002). Retrieved from http://www2.ed.gov/policy/elsec/leg/esea02/index.html

Obama, B. H. (2009, February 24). *Remarks of President Barack Obama: Address to joint session of Congress* [Transcript]. Retrieved from https://obamawhitehouse.archives.gov/video/EVR022409#transcript

Peat, M., Franklin, S., Devlin, M., & Charles, M. (2004, January). *Revisiting associations between student performance outcomes and formative assessment opportunities: Is there any impact on student learning?* Presentation at 21st ASCILITE Conference, Perth, Australia.

Peker, M., & Ertekin, E. (2011). The relationship between mathematics teaching anxiety and mathematics anxiety. *The New Educational Review*, *23*(1), 213–226.

Popham, W. J. (2008). *Classroom assessment: What teachers need to know* (5th ed.). Boston, MA: Allyn & Bacon.

Regional Educational Laboratory West. (2008). *Considerations in statewide implementation of computer adaptive testing*. Retrieved from http://relwest-archive.wested.org/system/memo_questions/11/attachments/original/Computer_20adaptive_20testing_20June_202008_1_.pdf

Riegle-Crumb, C. (2006). The path through math: Course sequences and academic performance at the intersection of race-ethnicity and gender. *American Journal of Education*, *113*(1), 101–122.

Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology*, *99*(3), 561–574.

Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*(3), 249–255.

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology, 95*(2), 240–257.

Roswal, G. M., Mims, A. A., Evans, M. D., Smith, B., Young, M., Burch, M., & Block, M. (1995). Effects of collaborative peer tutoring on urban seventh graders. *The Journal of Educational Research*, *88*(5), 275–279.

Rothman, T., & Henderson, M. (2011). Do school-based tutoring programs significantly improve student performance on standardized tests? *RMLE Online*, *34*(6), 1–10.

Schiller, K. S., Schmidt, W. H., Muller, C., & Houang, R. T. (2010). Hidden disparities: How courses and curricula shape opportunities in mathematics during high school. *Equity & Excellence in Education*, *43*(4), 414–433.

Schmidt, W. H., McKnight, C. C., Cogan, L. S., Jakwerth, P. M., & Houang, R. T. (1999). *Facing the consequences: Using TIMSS for a closer look at US mathematics and science education*. London, UK: Kluwer.

Schneider, M. (2010). *Finishing the first lap: The cost of first year student attrition in America's four-year colleges and universities*. Washington, DC: American Institutes for Research.

Shellard, E., & Moyer, P. S. (2002). *What principals need to know about teaching math*. Washington, DC: National Association of Elementary School Principals & Educational Research Service.

Slavin, R. E. (1999). How can funding equity ensure enhanced achievement? *Journal of Education Finance*, *24*(4), 519–528.

Sloan, T. (2010). A quantitative and qualitative study of math anxiety among preservice teachers. *The Educational Forum, 74*(3), 242–256.

Smith, T. M. (2014). *Instructional practices to support student success in Algebra 1*. Washington, DC: American Institutes of Research.

Snipes, J., & Finkelstein, N. (2015). *Opening a gateway to college access: Algebra at the right time.* Retrieved from ERIC database: https://files.eric.ed.gov/fulltext/ED559739.pdf

Snipes, J., Huang, C. W., Jaquet, K., & Finkelstein, N. (2016). *The effects of the Elevate Math summer program on math achievement and algebra readiness*. Washington, DC: Society for Research on Educational Effectiveness.

Stein, M. K., Kaufman, J. H., Sherman, M., & Hillen, A. F. (2011). Algebra: A challenge at the crossroads of policy and practice. *Review of Educational Research*, *81*(4), 453–492.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York, NY: Free Press.

Stone, C. (1998). Leveling the playing field: An urban school system examines equity in access to mathematics curriculum. *The Urban Review*, *30*(4), 295–307.

Strong American Schools. (2008). *Diploma to nowhere*. Retrieved from http://paworldclassmath.webs.com/8534051-Diploma-To-Nowhere-Strong-American-Schools-2008.pdf

Stronge, J. (2010). *Effective teachers = student achievement: What the research says*. New York, NY: Routledge.

Stronge, J. H. (2007). *Qualities of effective teachers*. Alexandria, VA: ASCD.

Stufflebeam, D. L. (2007). *CIPP evaluation model checklist: A tool for applying the CIPP Model to assess long-term enterprises.* Retrieved from University of Michigan website: https://www.wmich.edu/sites/default/files/attachments/u350/2014/cippchecklist_mar07.pdf

Suitts, S. (2015). *A new majority update: Low income students in the South and nation*. Atlanta, GA: Southern Education Foundation.

Telecsan, B. L., Slaton, D. B., & Stevens, K. B. (1999). Peer tutoring: Teaching students with learning disabilities to deliver time delay instruction. *Journal of Behavioral Education*, *9*(2), 133–154.

Triplett, C. F. (2004). Looking for a struggle: Exploring the emotions of a middle school reader. *Journal of Adolescent and Adult Literacy, 48*, 214–222.

Tseng, K. H., Diez, C. R., Lou, S. J., Tsai, H. L., & Tsai, T. S. (2010). Using the context, input, process, and product model to assess an engineering curriculum. *World Transactions on Engineering and Technology Education*, *8*(3), 256–259.

U.S. Department of Education. (1997). *Mathematics equals opportunity.* Retrieved from ERIC database: http://eric.ed.gov/?id=ED415119

Virginia Department of Education. (2012). *Algebra Readiness Diagnostic Test (ARDT): Tests, reports, and score interpretation information for teachers*. Retrieved from http://www.doe.virginia.gov/instruction/mathematics /middle/algebra_readiness/diagnostic_test/ardt_scoring_information.pdf

Virginia Department of Education. (2016a). *Statistics and reports: Enrollment.* Retrieved from http://www.doe.virginia.gov/statistics_reports/ enrollment/index.shtml

Virginia Department of Education. (2016b). Standard diploma: Minimum course and credit requirements. Retrieved from http://www.doe.virginia.gov/instruction/graduation/standard.shtml

Vilorio, D. (2014). STEM 101: Intro to tomorrow's jobs. *Occupational Outlook Quarterly*, *58*(1), 2–12.

Vogel, C. (2008). Algebra: Changing the equation. *District Administration*, *44*(6), 34–38.

Walker, S., & Senger, E. (2007). Using technology to teach developmental African-American algebra students. *The Journal of Computers in Mathematics and*

Science Teaching, *26*(3), 217–231.

Wallinger, L. (2008). *Improving the transition from middle grades to high schools*. Washington, DC: American Youth Policy Forum.

Walston, J., & McCarroll, J. C. (2010). *Eighth-grade algebra: Findings from the eighth-grade round of the Early Childhood Longitudinal Study, kindergarten class of 1998-99 statistics in brief* (NCES 2010-016). Washington, DC: National Center for Education Statistics.

Wasik, B. A. (1997). Volunteer tutoring programs. *Phi Delta Kappan*, *79*(4), 282–287.

Watson, J. (2008). *Blending learning: The convergence of online and face-to-face education*. Vienna, VA: North American Council for Online Learning.

Webb, M. (1988). Peer helping relationships in urban schools. *Equity and Choice*, *4*(3), 35–38.

Wiliam, D., & Thompson, M. (2007). Integrating assessment with learning: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping, teaching and learning* (pp. 53–82). Mahwah, NJ: Lawrence Erlbaum Associates.

World-class Instructional Design and Assessment. (2009). *Focus on Formative Assessment, 1*(2), 1–4.

Wu, S. S., Barth, M., Amin, H., Malcarne, V., & Menon, V. (2012). Math anxiety in second and third graders and its relation to mathematic achievement. *Frontiers in Psychology*, *3*(162).

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: Sage.

Zasky, J. (2016, July 19). The 5 most affordable cities in Virginia. Retrieved from http://livability.com/virginia/the-5-most-affordable-cities-in-virginia

Zuelke, D. C., & Nelson, J. G. (2001). The effect of a community agency's after-school tutoring program on reading and math GPA gains for at-risk tutored students. *Education*, *121*(4), 799–810.

VITA

Cassandra D. Harris

EDUCATION

Doctor of Education, College of William & Mary, Williamsburg, Virginia, May 2018.

Master of Science in Information Technologies and Business Management, Strayer University, Midlothian, Virginia, June 2001.

Bachelor of Science in Business Management, Saint Pauls College, Lawrenceville, Virginia, May 1998.

ACADEMIC EMPLOYMENT

Technical Support and Operations Manager, Department of Information Communication and Technology Services (ICTS), Richmond Public Schools, June 2001 – present.

Responsible for the conception and management of the service center servicing a district of 40+ school and administrative locations for call request care and resolution. The service center is the first line of support for all teacher, administrator and support staff for all areas of technology support such as student data retrieval, software/hardware integrations, and problem resolution. Additionally responsible for the technology supporting virtual servers, databases, security, telephones and network infrastructures for all student and staff data and communications. Assist in the Standards of Learning (SOL) and Algebra Readiness Diagnostic Test (ARDT) processes overseeing the training of technical talent for each administration of the assessment and providing support services. Serves as one of the district's project managers utilized for the software integration of student technology into the classroom.

PROFESSIONAL MEMBERSHIPS

American Society of Quality (ASQ)
International Society of Certified Electronics Technicians (ISCET)
Project Management Institute (PMI)